

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
26 January 2006 (26.01.2006)

PCT

(10) International Publication Number
WO 2006/010143 A2

(51) International Patent Classification:
A61K 38/26 (2006.01)

(21) International Application Number:
PCT/US2005/024781

(22) International Filing Date: 13 July 2005 (13.07.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/587,738 13 July 2004 (13.07.2004) US
60/608,723 10 September 2004 (10.09.2004) US

(71) Applicant (for all designated States except US): **NEOSE TECHNOLOGIES, INC.** [US/US]; 102 Witmer Road, Horsham, PA 19044 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **DEFREES, Shawn** [US/US]; 126 Filly Drive, North Wales, PA 19454 (US). **ZOPF, David, A.** [US/US]; 560 Beechtree Lane, Wayne, PA 19087 (US). **WANG, Zhi-Guang** [US/US]; 1473 Golden Drive, Dresher, PA 19015 (US). **CLAUSEN, Henrik** [DK/DK]; Soebrinken 6, DK-2840 Holte (DK).

(74) Agents: **ESKER, Todd** et al.; Morgan Lewis & Bockius LLP, 2 Palo Alto Square, Suite 900, 3000 El Camino Real, Palo Alto, CA 94304 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: BRANCHED PEG REMODELING AND GLYCOSYLATION OF GLUCAGON-LIKE PEPTIDE-1 [GLP-1]

(57) Abstract: The present invention provides polypeptides that include an O-linked glycoconjugate in which a species such as a water-soluble polymer, a therapeutic agent of a biomolecule is covalently linked through an intact O-linked glycosyl residue to the polypeptide. The polypeptides of the invention include wild-type peptides and mutant peptides that include an O-linked glycosylation site that is not present in the wild-type peptide. Also provided are methods of making the peptides of the invention and methods, pharmaceutical compositions containing the peptides and methods of treating, ameliorating or preventing disease in mammals by administering an amount of a peptide of the invention sufficient to achieve the desired response.



WO 2006/010143 A2

BRANCHED PEG REMODELING AND GLYCOSYLATION OF GLUCAGON-LIKE PEPTIDE-1 [GLP-1]

Field of the Invention

5 [0001] The present invention relates to O-linked glycosylated glycopeptides, particularly glucagon-like peptide-1 (GLP-1) and GLP-1 peptide mutants that include O-linked glycosylation sites not present in the wild-type peptide.

BACKGROUND OF THE INVENTION

10 [0002] Glucagon-like peptide-1 (GLP-1) is an important glucorecretin hormone secreted from intestinal L cells in response to nutrient ingestion. GLP-1 functions to regulate plasma glucose levels via various independent mechanisms, making it an ideal candidate for treatment of diabetes, and possibly useful in the pharmacotherapy of obesity.

15 [0003] The biologically active forms of GLP-1 possess multiple functions *in vivo*, including enhancement of glucose-dependent insulin secretion, stimulation of proinsulin gene expression, and suppression of glucagon secretion and gastric emptying. GLP-1 also enhances insulin sensitivity, induces β cell differentiation and proliferation, decreases caloric intake, and increases satiety.

20 [0004] The mature, active form of GLP-1 is a 30 amino acid derivative of proglucagon, a 160 amino acid prohormone. GLP-1 is synthesized by post translational processing of proglucagon in intestinal L cells. Posttranslational processing of proglucagon gives rise to glucagon, GLP-1, GLP-2 and other peptide sequences, IP-1 and IP-2, of unknown function. The initial GLP-1 cleaved from proglucagon is further processed first by N-terminal cleavage to form a biologically active peptide (GLP-1₍₇₋₃₇₎). GLP-1₍₇₋₃₇₎ is then C-terminally truncated and amidated to form the predominant biologically active species, GLP-1₍₇₋₃₆₎amide.

25 [0005] GLP-1₍₇₋₃₆₎amide has a very short half life *in vivo*. The plasma half life of GLP-1 is about 5 minutes, and the metabolic clearance rate is about 12-13 minutes. In circulation, the predominant form of GLP-1 is rapidly inactivated as a result of degradation by dipeptidyl-peptidase IV (see *e.g.*, Deacon *et al.* (1995) *Endocrinol. Metab.* 80:952-957, and Hansen *et al.* (1999) *Endocrinology* 140:5356). GLP-1₍₇₋₃₆₎amide is also susceptible to degradation by
30 neutral endopeptidases, including NEP 24.11 (Sodman *et al.* (1995) *Reg. Peptides* 58:149-156).

[0006] The unique ability of GLP-1 to lower postprandial hyperglycemia via three independent and complementary mechanisms of action (increased insulin secretion, inhibition of glucagon release, and inhibition of gastrointestinal motility) are what make this peptide hormone an ideal candidate for the treatment of diabetes. Indeed, GLP-1 provides
5 unprecedented advantages over any other pharmacological agent currently available. Unfortunately, despite its potential, there are serious limitations to the possible therapeutic use of GLP-1 in humans. The most serious limitation is the very short half life of GLP-1 *in vivo*. Even when administered subcutaneously, peak concentrations return to baseline within 90 minutes.

10 [0007] The therapeutic potential of GLP-1 and its very short half life have prompted the search for and discovery of analogs that may provide an extended GLP-1-like biological activity. Several analogs have been isolated from other species (Fehmann, H.C., *et al.* (1995) *Endocrine Reviews* 16:390-410, and Thorens B. *et al.* (1993) *Diabetes* 42:1678-1682), and mutant GLP-1 peptides resistant to degradation have been created (Xiao *et al.* (2001)
15 *Biochemistry* 40:2860-2869).

[0008] Some GLP-1 analogs may show some promise as therapeutics. However, since GLP-1 peptide is a highly multifunctional protein, mutants and interspecies homologs may have unpredictable pleiotropic effects. Indeed, Xiao *et al.* showed that some mutants exhibit altered biological activity independent of any changes in receptor binding activity. Thus, the
20 biological activities of GLP-1 can be uncoupled from one another.

[0009] Diabetes, obesity and other disorders of sugar metabolism and glycemic control carry a very high price for the individual, as well as the society in terms of health, lost productivity and the loss of wages and financial output. Thus, there is clearly a need in the art for effective medications that facilitate glycemic control in the individual. A stabilized
25 GLP-1 with increased half life *in vivo* could meet this need. Preferably a stabilized GLP-1 peptide would be very similar to the wild type protein, such that changes to biological activity, and hence possible side effects of therapy can be minimized. The present invention answers the need for stabilized GLP-1 molecules, thereby providing therapeutically effective GLP-1 peptides. Other objects and advantages will become apparent from the detailed
30 description that follows.

BRIEF SUMMARY OF THE INVENTION

[0010] Diabetes and disorders of glycemic control are serious conditions which, if unchecked can have dire consequences for the individual and society at large. Although type 1 diabetes can be controlled more or less effectively with insulin injections, there are multiple pathways of glycemic control. If some of those pathway could also be recruited into therapeutic methods, glycemic control for diabetics would be improved. Further, enhanced glycemic control for type 2 diabetics and individuals struggling with obesity, could provide enhanced health benefits for these groups of individuals as well.

[0011] As noted above, Glucagon-Like Peptide-1 (GLP-1) facilitates glycemic control in the individual by multiple mechanisms. Thus, GLP-1 is an ideal candidate for the pharmacotherapy of glycemic disorders. Unfortunately, the potential therapeutic uses of GLP-1 are limited by the short *in vivo* half life of the protein. Fortunately, methods that improve *in vivo* half life of the protein have now been discovered. These methods have the added advantage that they introduce minimal alterations to the protein and therefore the risks of side effects are minimized.

[0012] Indeed, it has now been discovered that enzymatic glycoconjugation reactions can be specifically targeted to O-linked glycosylation sites and to glycosyl residues that are attached to O-linked glycosylation sites. The targeted O-linked glycosylation sites can be sites native to a wild-type peptide or, alternatively, they can be introduced into a peptide by mutation. Thus, a method for prolonging the *in vivo* half life of GLP-1 (and other proteins) is provided by the methods of the invention.

[0013] In addition to the discovery that O-linked glycosylation sites, and glycosyl residues linked thereto, are useful targets for glycoconjugation reactions, the present invention provides mutant polypeptides in which the amino acid sequence is manipulated by mutation to insert, remove or relocate one or more O-linked glycosylation site in the peptide. When a site is added or relocated, it is not present or not present in a selected location in the wild type peptide. The mutant O-linked glycosylation site is a point of attachment for a modified glycosyl residue that is enzymatically conjugated to the O-linked glycosylation site. Using the methods of the invention, the glycosylation site can be shifted to any efficacious position on the peptide. For example, if the native glycosylation site is sufficiently proximate the active site of the peptide that conjugation of a large water-soluble polymer interferes with the biological activity of the peptide, it is within the scope of the invention to engineer a mutant

peptide that includes an O-linked glycosylation site as removed from the active site as necessary to provide a biologically active peptide conjugate.

[0014] Post-expression *in vitro* modification of peptides is an attractive strategy to remedy the deficiencies of methods that rely on controlling glycosylation by engineering expression systems; including both modification of glycan structures or introduction of glycans at novel sites. A comprehensive toolbox of recombinant eukaryotic glycosyltransferases is becoming available, making *in vitro* enzymatic synthesis of mammalian glycoconjugates with custom designed glycosylation patterns and glycosyl structures possible. See, for example, U.S. Patent No. 5,876,980; 6,030,815; 5,728,554; 5,922,577; and WO/9831826; US2003180835; and WO 03/031464.

[0015] *In vitro* glycosylation offers a number of advantages compared to recombinant expression of glycoproteins of which custom design and higher degree of homogeneity of the glycosyl moiety are examples. Moreover, combining bacterial expression of glycotherapeutics with *in vitro* modification (or placement) of the glycosyl residue offers numerous advantages over traditional recombinant expression technology including reduced potential exposure to adventitious agents, increased homogeneity of product, and cost reduction.

[0016] In addition to methods of O-linked glycosylation, inserting O-linked glycosylation sites into peptides and methods of glycosylating the inserted sites, the present invention provides methods of improving pharmacological parameters of glycopeptide therapeutics, e.g., altering pharmacokinetics, pharmacodynamics and bioavailability of therapeutic (glyco)proteins, e.g., hormones, and enzymes. In particular, the invention provides a method for lengthening the *in vivo* half-lives of glycopeptide therapeutics by conjugating a water-soluble polymer to the peptide through an intact glycosyl linking group. In an exemplary embodiment, covalent attachment of polymers, such as polyethylene glycol (PEG), e.g., m-PEG, to such peptides affords conjugates having *in vivo* residence times, and pharmacokinetic and pharmacodynamic properties, enhanced relative to the unconjugated peptide.

[0017] Art-recognized methods of covalent PEGylation rely on chemical conjugation through reactive groups on amino acids or carbohydrates. A major shortcoming of chemical conjugation of PEG to proteins or glycoproteins is lack of selectivity, which often results in attachment of PEG at sites implicated in protein or glycoprotein bioactivity. Several

strategies have been developed to address site selective conjugation chemistries, however, one universal method suitable for a variety of recombinant proteins has yet to be developed.

[0018] In contrast to art-recognized methods, the present invention provides a novel strategy for highly selective site directed O-linked glycoconjugation, *e.g.*, glyco-PEGylation.

5 In an exemplary embodiment of the invention, site directed attachment sites for PEGylation are provided by *in vitro* enzymatic GalNAc O-linked glycosylation of specific peptide sequences, *e.g.*, mutant sequences, containing serine and threonine residues. The recombinant proteins are preferably expressed in bacteria, *e.g.*, *E. coli*, to avoid host cell glycosylation. Glyco-PEGylation is subsequently performed enzymatically utilizing a
10 glycosyltransferase, *e.g.*, a sialyltransferase, capable of transferring the species PEG-glycosyl, *e.g.*, PEG-sialic acid, to an O-linked glycosylation site ("glyco-PEGylation"). O-linked glycosylation sites may be introduced into any peptide sequence by providing a mutant peptide with simple short sequence motifs.

[0019] Additional aspects, advantages and objects of the present invention will be apparent
15 from the detailed description that follows.

BRIEF DESCRIPTION OF THE FIGURES

[0020] FIG. 1 provides a list of peptides which can be used as substrates for this invention.

DETAILED DESCRIPTION OF THE INVENTION

Abbreviations

20 [0021] PEG, poly(ethyleneglycol); m-PEG, methoxy-poly(ethylene glycol); PPG, poly(propyleneglycol); m-PPG, methoxy-poly(propylene glycol); Fuc, fucosyl; Gal, galactosyl; GalNAc, N-acetylgalactosaminyl; Glc, glucosyl; GlcNAc, N-acetylglucosaminyl; Man, mannosyl; ManAc, mannosaminyl acetate; Sia, sialic acid; and NeuAc, N-acetylneuraminyl.

Definitions

25 [0022] Unless defined otherwise, all technical and scientific terms used herein generally have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Generally, the nomenclature used herein and the laboratory
30 procedures in cell culture, molecular genetics, organic chemistry and nucleic acid chemistry and hybridization are those well known and commonly employed in the art. Standard techniques are used for nucleic acid and peptide synthesis. The techniques and procedures

are generally performed according to conventional methods in the art and various general references (*see generally*, Sambrook *et al.* MOLECULAR CLONING: A LABORATORY MANUAL, 2d ed. (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., which is incorporated herein by reference), which are provided throughout this document. The nomenclature used herein and the laboratory procedures in analytical chemistry, and organic synthetic described below are those well known and commonly employed in the art. Standard techniques, or modifications thereof, are used for chemical syntheses and chemical analyses.

[0023] The term “nucleic acid” or “polynucleotide” refers to deoxyribonucleic acids (DNA) or ribonucleic acids (RNA) and polymers thereof in either single- or double-stranded form. Unless specifically limited, the term encompasses nucleic acids containing known analogues of natural nucleotides that have similar binding properties as the reference nucleic acid and are metabolized in a manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (*e.g.*, degenerate codon substitutions), alleles, orthologs, SNPs, and complementary sequences as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer *et al.*, *Nucleic Acid Res.* **19**:5081 (1991); Ohtsuka *et al.*, *J. Biol. Chem.* **260**:2605-2608 (1985); and Rossolini *et al.*, *Mol. Cell. Probes* **8**:91-98 (1994)). The term nucleic acid is used interchangeably with gene, cDNA, and mRNA encoded by a gene.

[0024] The term “gene” means the segment of DNA involved in producing a polypeptide chain. It may include regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons).

[0025] The term “isolated,” when applied to a nucleic acid or protein, denotes that the nucleic acid or protein is essentially free of other cellular components with which it is associated in the natural state. It is preferably in a homogeneous state although it can be in either a dry or aqueous solution. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein that is the predominant species present in a preparation is substantially purified. In particular, an isolated gene is separated from open reading frames that flank the gene and encode a protein other than the gene of interest. The

term “purified” denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Particularly, it means that the nucleic acid or protein is at least 85% pure, more preferably at least 95% pure, and most preferably at least 99% pure.

[0026] The term “amino acid” refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, *e.g.*, hydroxyproline, γ -carboxyglutamate, and O-phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, *i.e.*, an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, *e.g.*, homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (*e.g.*, norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. “Amino acid mimetics” refers to chemical compounds having a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

[0027] There are various known methods in the art that permit the incorporation of an unnatural amino acid derivative or analog into a polypeptide chain in a site-specific manner, *see, e.g.*, WO 02/086075.

[0028] Amino acids may be referred to herein by either the commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

[0029] “Conservatively modified variants” applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, “conservatively modified variants” refers to those nucleic acids that encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are “silent variations,” which are one species of conservatively modified variations. Every nucleic acid

sequence herein that encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule.

5 Accordingly, each silent variation of a nucleic acid that encodes a polypeptide is implicit in each described sequence.

[0030] As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded
10 sequence is a "conservatively modified variant" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

15 [0031] The following eight groups each contain amino acids that are conservative substitutions for one another:

- 1) Alanine (A), Glycine (G);
- 2) Aspartic acid (D), Glutamic acid (E);
- 3) Asparagine (N), Glutamine (Q);
- 20 4) Arginine (R), Lysine (K);
- 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V);
- 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W);
- 7) Serine (S), Threonine (T); and
- 8) Cysteine (C), Methionine (M)

25 (*see, e.g.,* Creighton, *Proteins* (1984)).

[0032] Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

30 [0033] In the present application, amino acid residues are numbered according to their relative positions from the left most residue, which is numbered 1, in an unmodified wild-type polypeptide sequence.

[0034] "Peptide" refers to a polymer in which the monomers are amino acids and are joined together through amide bonds, alternatively referred to as a polypeptide. Additionally, unnatural amino acids, for example, β -alanine, phenylglycine and homoarginine are also included. Amino acids that are not gene-encoded may also be used in the present invention. Furthermore, amino acids that have been modified to include reactive groups, glycosylation sites, polymers, therapeutic moieties, biomolecules and the like may also be used in the invention. All of the amino acids used in the present invention may be either the D - or L - isomer. The L -isomer is generally preferred. In addition, other peptidomimetics are also useful in the present invention. As used herein, "peptide" refers to both glycosylated and unglycosylated peptides. Also included are peptides that are incompletely glycosylated by a system that expresses the peptide. For a general review, *see*, Spatola, A. F., in CHEMISTRY AND BIOCHEMISTRY OF AMINO ACIDS, PEPTIDES AND PROTEINS, B. Weinstein, eds., Marcel Dekker, New York, p. 267 (1983).

[0035] In the present application, amino acid residues are numbered according to their relative positions from the left most residue, which is numbered 1, in a peptide sequence.

[0036] "Proximate a proline residue," as used herein refers to an amino acid that is less than about 10 amino acids removed from a proline residue, preferably, less than about 9, 8, 7, 6 or 5 amino acids removed from a proline residue, more preferably, less than about 4, 3, 2 or 1 residues removed from a proline residue. The amino acid "proximate a proline residue" may be on the C- or N-terminal side of the proline residue.

[0037] The term "sialic acid" refers to any member of a family of nine-carbon carboxylated sugars. The most common member of the sialic acid family is N-acetyl-neuraminic acid (2-keto-5-acetamido-3,5-dideoxy-D-glycero-D-galactononulopyranos-1-onic acid (often abbreviated as Neu5Ac, NeuAc, or NANA). A second member of the family is N-glycolyl-neuraminic acid (Neu5Gc or NeuGc), in which the N-acetyl group of NeuAc is hydroxylated. A third sialic acid family member is 2-keto-3-deoxy-nonulosonic acid (KDN) (Nadano *et al.* (1986) *J. Biol. Chem.* **261**: 11550-11557; Kanamori *et al.*, *J. Biol. Chem.* **265**: 21811-21819 (1990)). Also included are 9-substituted sialic acids such as a 9-O-C₁-C₆ acyl-Neu5Ac like 9-O-lactyl-Neu5Ac or 9-O-acetyl-Neu5Ac, 9-deoxy-9-fluoro-Neu5Ac and 9-azido-9-deoxy-Neu5Ac. For review of the sialic acid family, *see, e.g.*, Varki, *Glycobiology* **2**: 25-40 (1992); *Sialic Acids: Chemistry, Metabolism and Function*, R. Schauer, Ed. (Springer-Verlag, New

York (1992)). The synthesis and use of sialic acid compounds in a sialylation procedure is disclosed in international application WO 92/16640, published October 1, 1992.

[0038] As used herein, the term “modified sugar,” refers to a naturally- or non-naturally-occurring carbohydrate that is enzymatically added onto an amino acid or a glycosyl residue of a peptide in a process of the invention. The modified sugar is selected from a number of enzyme substrates including, but not limited to sugar nucleotides (mono-, di-, and tri-phosphates), activated sugars (*e.g.*, glycosyl halides, glycosyl mesylates) and sugars that are neither activated nor nucleotides. The “modified sugar” is covalently functionalized with a “modifying group.” Useful modifying groups include, but are not limited to, water-soluble polymers, therapeutic moieties, diagnostic moieties, biomolecules and the like. The modifying group is preferably not a naturally occurring, or an unmodified carbohydrate. The locus of functionalization with the modifying group is selected such that it does not prevent the “modified sugar” from being added enzymatically to a peptide.

[0039] The term “water-soluble” refers to moieties that have some detectable degree of solubility in water. Methods to detect and/or quantify water solubility are well known in the art. Exemplary water-soluble polymers include peptides, saccharides, poly(ethers), poly(amines), poly(carboxylic acids) and the like. Peptides can have mixed sequences of be composed of a single amino acid, *e.g.*, poly(lysine). An exemplary polysaccharide is poly(sialic acid). An exemplary poly(ether) is poly(ethylene glycol), *e.g.*, m-PEG. Poly(ethylene imine) is an exemplary polyamine, and poly(acrylic) acid is a representative poly(carboxylic acid).

[0040] The polymer backbone of the water-soluble polymer can be poly(ethylene glycol) (*i.e.* PEG). However, it should be understood that other related polymers are also suitable for use in the practice of this invention and that the use of the term PEG or poly(ethylene glycol) is intended to be inclusive and not exclusive in this respect. The term PEG includes poly(ethylene glycol) in any of its forms, including alkoxy PEG, difunctional PEG, multiarmed PEG, forked PEG, branched PEG, pendent PEG (*i.e.* PEG or related polymers having one or more functional groups pendent to the polymer backbone), or PEG with degradable linkages therein.

[0041] The polymer backbone can be linear or branched. Branched polymer backbones are generally known in the art. Typically, a branched polymer has a central branch core moiety and a plurality of linear polymer chains linked to the central branch core. PEG is commonly

used in branched forms that can be prepared by addition of ethylene oxide to various polyols, such as glycerol, pentaerythritol and sorbitol. The central branch moiety can also be derived from several amino acids, such as lysine. The branched poly(ethylene glycol) can be represented in general form as R(-PEG-OH)_m in which R represents the core moiety, such as glycerol or pentaerythritol, and m represents the number of arms. Multi-armed PEG molecules, such as those described in U.S. Pat. No. 5,932,462, which is incorporated by reference herein in its entirety, can also be used as the polymer backbone.

[0042] Many other polymers are also suitable for the invention. Polymer backbones that are non-peptidic and water-soluble, with from 2 to about 300 termini, are particularly useful in the invention. Examples of suitable polymers include, but are not limited to, other poly(alkylene glycols), such as poly(propylene glycol) ("PPG"), copolymers of ethylene glycol and propylene glycol and the like, poly(oxyethylated polyol), poly(olefinic alcohol), poly(vinylpyrrolidone), poly(hydroxypropylmethacrylamide), poly(α -hydroxy acid), poly(vinyl alcohol), polyphosphazene, polyoxazoline, poly(N-acryloylmorpholine), such as described in U.S. Pat. No. 5,629,384, which is incorporated by reference herein in its entirety, and copolymers, terpolymers, and mixtures thereof. Although the molecular weight of each chain of the polymer backbone can vary, it is typically in the range of from about 100 Da to about 100,000 Da, often from about 6,000 Da to about 80,000 Da.

[0043] The term "glycoconjugation," as used herein, refers to the enzymatically mediated conjugation of a modified sugar species to an amino acid or glycosyl residue of a polypeptide, *e.g.*, a mutant human growth hormone of the present invention. A subgenus of "glycoconjugation" is "glycol-PEGylation," in which the modifying group of the modified sugar is poly(ethylene glycol), and alkyl derivative (*e.g.*, m-PEG) or reactive derivative (*e.g.*, H₂N-PEG, HOOC-PEG) thereof.

[0044] The terms "large-scale" and "industrial-scale" are used interchangeably and refer to a reaction cycle that produces at least about 250 mg, preferably at least about 500 mg, and more preferably at least about 1 gram of glycoconjugate at the completion of a single reaction cycle.

[0045] The term, "glycosyl linking group," as used herein refers to a glycosyl residue to which a modifying group (*e.g.*, PEG moiety, therapeutic moiety, biomolecule) is covalently attached; the glycosyl linking group joins the modifying group to the remainder of the conjugate. In the methods of the invention, the "glycosyl linking group" becomes covalently

attached to a glycosylated or unglycosylated peptide, thereby linking the agent to an amino acid and/or glycosyl residue on the peptide. A “glycosyl linking group” is generally derived from a “modified sugar” by the enzymatic attachment of the “modified sugar” to an amino acid and/or glycosyl residue of the peptide. The glycosyl linking group can be a saccharide-derived structure that is degraded during formation of modifying group-modified sugar cassette (e.g., oxidation→Schiff base formation→reduction), or the glycosyl linking group may be intact. An “intact glycosyl linking group” refers to a linking group that is derived from a glycosyl moiety in which the saccharide monomer that links the modifying group and to the remainder of the conjugate is not degraded, e.g., oxidized, e.g., by sodium metaperiodate. “Intact glycosyl linking groups” of the invention may be derived from a naturally occurring oligosaccharide by addition of glycosyl unit(s) or removal of one or more glycosyl unit from a parent saccharide structure.

[0046] The term “targeting moiety,” as used herein, refers to species that will selectively localize in a particular tissue or region of the body. The localization is mediated by specific recognition of molecular determinants, molecular size of the targeting agent or conjugate, ionic interactions, hydrophobic interactions and the like. Other mechanisms of targeting an agent to a particular tissue or region are known to those of skill in the art. Exemplary targeting moieties include antibodies, antibody fragments, transferrin, HS-glycoprotein, coagulation factors, serum proteins, β -glycoprotein, G-CSF, GM-CSF, M-CSF, EPO and the like.

[0047] As used herein, “therapeutic moiety” means any agent useful for therapy including, but not limited to, antibiotics, anti-inflammatory agents, anti-tumor drugs, cytotoxins, and radioactive agents. “Therapeutic moiety” includes prodrugs of bioactive agents, constructs in which more than one therapeutic moiety is bound to a carrier, e.g., multivalent agents.

Therapeutic moiety also includes proteins and constructs that include proteins. Exemplary proteins include, but are not limited to, Glucagon like protein-1 (GLP-1), Erythropoietin (EPO), Granulocyte Colony Stimulating Factor (GCSF), Granulocyte Macrophage Colony Stimulating Factor (GMCSF), Interferon (e.g., Interferon- α , - β , - γ), Interleukin (e.g., Interleukin II), serum proteins (e.g., Factors VII, VIIa, VIII, IX, and X), Human Chorionic Gonadotropin (HCG), Follicle Stimulating Hormone (FSH) and Lutenizing Hormone (LH) and antibody fusion proteins (e.g. Tumor Necrosis Factor Receptor ((TNFR)/Fc domain fusion protein)).

[0048] As used herein, "anti-tumor drug" means any agent useful to combat cancer including, but not limited to, cytotoxins and agents such as antimetabolites, alkylating agents, anthracyclines, antibiotics, antimitotic agents, procarbazine, hydroxyurea, asparaginase, corticosteroids, interferons and radioactive agents. Also encompassed within the scope of the term "anti-tumor drug," are conjugates of peptides with anti-tumor activity, *e.g.* TNF- α . Conjugates include, but are not limited to those formed between a therapeutic protein and a glycoprotein of the invention. A representative conjugate is that formed between PSGL-1 and TNF- α .

[0049] As used herein, "a cytotoxin or cytotoxic agent" means any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, teniposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracenedione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Other toxins include, for example, ricin, CC-1065 and analogues, the duocarmycins. Still other toxins include diphtheria toxin, and snake venom (*e.g.*, cobra venom).

[0050] As used herein, "a radioactive agent" includes any radioisotope that is effective in diagnosing or destroying a tumor. Examples include, but are not limited to, indium-111, cobalt-60. Additionally, naturally occurring radioactive elements such as uranium, radium, and thorium, which typically represent mixtures of radioisotopes, are suitable examples of a radioactive agent. The metal ions are typically chelated with an organic chelating moiety.

[0051] Many useful chelating groups, crown ethers, cryptands and the like are known in the art and can be incorporated into the compounds of the invention (*e.g.*, EDTA, DTPA, DOTA, NTA, HDTA, *etc.* and their phosphonate analogs such as DTPP, EDTP, HDTP, NTP, *etc.*).

See, for example, Pitt *et al.*, "The Design of Chelating Agents for the Treatment of Iron Overload," In, INORGANIC CHEMISTRY IN BIOLOGY AND MEDICINE; Martell, Ed.; American Chemical Society, Washington, D.C., 1980, pp. 279-312; Lindoy, THE CHEMISTRY OF MACROCYCLIC LIGAND COMPLEXES; Cambridge University Press, Cambridge, 1989; Dugas, BIOORGANIC CHEMISTRY; Springer-Verlag, New York, 1989, and references contained therein.

[0052] Additionally, a manifold of routes allowing the attachment of chelating agents, crown ethers and cyclodextrins to other molecules is available to those of skill in the art. *See*,

for example, Meares *et al.*, "Properties of In Vivo Chelate-Tagged Proteins and Polypeptides." In, MODIFICATION OF PROTEINS: FOOD, NUTRITIONAL, AND PHARMACOLOGICAL ASPECTS;" Feeney, *et al.*, Eds., American Chemical Society, Washington, D.C., 1982, pp. 370-387; Kasina *et al.*, *Bioconjugate Chem.*, **9**: 108-117 (1998);
5 Song *et al.*, *Bioconjugate Chem.*, **8**: 249-255 (1997).

[0053] As used herein, "pharmaceutically acceptable carrier" includes any material, which when combined with the conjugate retains the conjugates' activity and is non-reactive with the subject's immune systems. Examples include, but are not limited to, any of the standard pharmaceutical carriers such as a phosphate buffered saline solution, water, emulsions such
10 as oil/water emulsion, and various types of wetting agents. Other carriers may also include sterile solutions, tablets including coated tablets and capsules. Typically such carriers contain excipients such as starch, milk, sugar, certain types of clay, gelatin, stearic acid or salts thereof, magnesium or calcium stearate, talc, vegetable fats or oils, gums, glycols, or other known excipients. Such carriers may also include flavor and color additives or other
15 ingredients. Compositions comprising such carriers are formulated by well known conventional methods.

[0054] As used herein, "administering" means oral administration, administration as a suppository, topical contact, intravenous, intraperitoneal, intramuscular, intralesional, or subcutaneous administration, administration by inhalation, or the implantation of a slow-
20 release device, *e.g.*, a mini-osmotic pump, to the subject. Administration is by any route including parenteral and transmucosal (*e.g.*, oral, nasal, vaginal, rectal, or transdermal), particularly by inhalation. Parenteral administration includes, *e.g.*, intravenous, intramuscular, intra-arteriole, intradermal, subcutaneous, intraperitoneal, intraventricular, and intracranial. Moreover, where injection is to treat a tumor, *e.g.*, induce apoptosis,
25 administration may be directly to the tumor and/or into tissues surrounding the tumor. Other modes of delivery include, but are not limited to, the use of liposomal formulations, intravenous infusion, transdermal patches, etc.

[0055] The term "isolated" refers to a material that is substantially or essentially free from components, which are used to produce the material. For peptide conjugates of the invention,
30 the term "isolated" refers to material that is substantially or essentially free from components, which normally accompany the material in the mixture used to prepare the peptide conjugate. "Isolated" and "pure" are used interchangeably. Typically, isolated peptide conjugates of the

invention have a level of purity preferably expressed as a range. The lower end of the range of purity for the peptide conjugates is about 60%, about 70% or about 80% and the upper end of the range of purity is about 70%, about 80%, about 90% or more than about 90%.

[0056] When the peptide conjugates are more than about 90% pure, their purities are also preferably expressed as a range. The lower end of the range of purity is about 90%, about 92%, about 94%, about 96% or about 98%. The upper end of the range of purity is about 92%, about 94%, about 96%, about 98% or about 100% purity.

[0057] Purity is determined by any art-recognized method of analysis (*e.g.*, band intensity on a silver stained gel, polyacrylamide gel electrophoresis, HPLC, or a similar means).

[0058] “Essentially each member of the population,” as used herein, describes a characteristic of a population of peptide conjugates of the invention in which a selected percentage of the modified sugars added to a peptide are added to multiple, identical acceptor sites on the peptide. “Essentially each member of the population” speaks to the “homogeneity” of the sites on the peptide conjugated to a modified sugar and refers to conjugates of the invention, which are at least about 80%, preferably at least about 90% and more preferably at least about 95% homogenous.

[0059] “Homogeneity,” refers to the structural consistency across a population of acceptor moieties to which the modified sugars are conjugated. Thus, in a peptide conjugate of the invention in which each modified sugar moiety is conjugated to an acceptor site having the same structure as the acceptor site to which every other modified sugar is conjugated, the peptide conjugate is said to be about 100% homogeneous. Homogeneity is typically expressed as a range. The lower end of the range of homogeneity for the peptide conjugates is about 60%, about 70% or about 80% and the upper end of the range of purity is about 70%, about 80%, about 90% or more than about 90%.

[0060] When the peptide conjugates are more than or equal to about 90% homogeneous, their homogeneity is also preferably expressed as a range. The lower end of the range of homogeneity is about 90%, about 92%, about 94%, about 96% or about 98%. The upper end of the range of purity is about 92%, about 94%, about 96%, about 98% or about 100% homogeneity. The purity of the peptide conjugates is typically determined by one or more methods known to those of skill in the art, *e.g.*, liquid chromatography-mass spectrometry (LC-MS), matrix assisted laser desorption mass time of flight spectrometry (MALDITOF), capillary electrophoresis, and the like.

[0061] “Substantially uniform glycoform” or a “substantially uniform glycosylation pattern,” when referring to a glycopeptide species, refers to the percentage of acceptor moieties that are glycosylated by the glycosyltransferase of interest (*e.g.*, fucosyltransferase). For example, in the case of a α 1,2 fucosyltransferase, a substantially uniform fucosylation pattern exists if substantially all (as defined below) of the Gal β 1,4-GlcNAc-R and sialylated analogues thereof are fucosylated in a peptide conjugate of the invention. It will be understood by one of skill in the art, that the starting material may contain glycosylated acceptor moieties (*e.g.*, fucosylated Gal β 1,4-GlcNAc-R moieties). Thus, the calculated percent glycosylation will include acceptor moieties that are glycosylated by the methods of the invention, as well as those acceptor moieties already glycosylated in the starting material.

[0062] The term “substantially” in the above definitions of “substantially uniform” generally means at least about 40%, at least about 70%, at least about 80%, or more preferably at least about 90%, and still more preferably at least about 95% of the acceptor moieties for a particular glycosyltransferase are glycosylated.

Introduction

[0063] The present invention provides stabilized peptides for therapeutic use. In one embodiment the invention provides conjugates of glycopeptides in which a modified sugar moiety is attached either directly or indirectly (*e.g.*, through and intervening glycosyl residue) to an O-linked glycosylation site on the peptide. Also provided are methods for producing the conjugates of the invention.

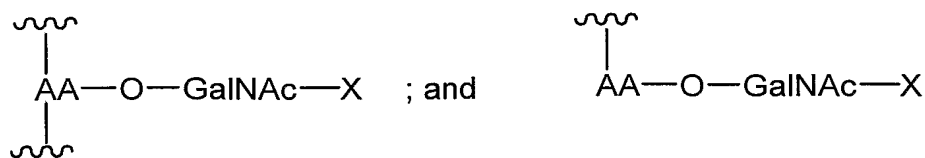
[0064] The O-linked glycosylation site is generally the hydroxy side chain of a natural (*e.g.*, serine, threonine) or unnatural (*e.g.*, 5-hydroxyproline or 5-hydroxylysine) amino acid. Exemplary O-linked saccharyl residues include N-acetylgalactosamine, galactose, mannose, GlcNAc, glucose, fucose or xylose.

[0065] The methods of the invention can be practiced on any peptide having an O-linked glycosylation site. For example, in some embodiments the methods are of use to produce O-linked glycoconjugates in which the glycosyl moiety is attached to an O-linked glycosylation site that is present in the wild type peptide.

[0066] In other embodiments the invention provides novel mutant peptides that include one or more O-linked glycosylation site(s) that is/are not present in the wild-type peptide. Also provided are O-linked glycosylated versions of the mutant peptides, and methods of

preparing O-linked glycosylated mutant peptides. Additional methods include the elaboration, trimming back and/or modification of the O-linked glycosyl residue and glycosyl residues that are N-, rather than O-linked.

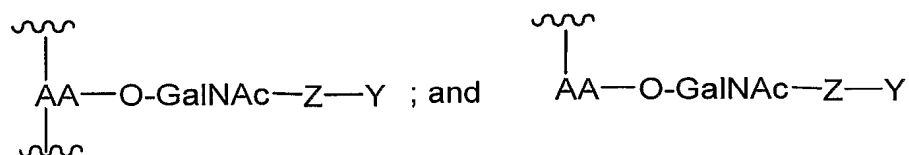
[0067] In an exemplary aspect, the invention provides a mutant peptide having the formula:



in which AA is an amino acid with a side chain that includes a hydroxyl moiety. Exemplary hydroxyamino acids are threonine and serine. The GalNAc moiety is linked to AA through the oxygen atom of the hydroxyl moiety. AA may be present in the wild type peptide or, alternatively, it is added or relocated by mutating the sequence of the wild type peptide. X is a modifying group or it is a saccharyl moiety, *e.g.*, sialyl, galactosyl and Gal-Sia groups. In an exemplary embodiment, in which X is a saccharyl moiety, it includes a modifying group, as discussed herein.

[0068] As shown in the formulae above, the glycosylated amino acid can be at the N- or C-peptide terminus or internal to the peptide sequence.

[0069] In another exemplary embodiment, the invention provides a peptide conjugate having the formula:



in which Z is a bond or a saccharyl residue selected from Gal, Sia and Gal-Sia. Y is a modifying group. The saccharyl residue bearing the modifying group (“glycosyl linking group”) is enzymatically attached to the peptide-tethered glycosyl residue, *e.g.*, forming an intact glycosyl linking group between the modified sugar and the remainder of the peptide-tethered glycosyl residue.

[0070] In yet another exemplary embodiment, AA is located within a proline-rich segment of the mutant peptide and/or it is proximate a proline residue. Appropriate sequences forming O-linked glycosylation sites are readily determined by interrogating the enzymatic O-linked glycosylation of short peptides containing one or more putative O-linked

glycosylation sites. In another exemplary embodiment, O-linked glycosylation sites can be created at any position in a molecule, using techniques well known in the art. Peptides with introduced O-linked glycosylation sites can be tested for biological activity according to the methods of the invention.

5 [0071] The conjugates of the invention are formed between peptides and diverse species such as water-soluble polymers, therapeutic moieties, diagnostic moieties, targeting moieties and the like. Also provided are conjugates that include two or more peptides linked together through a linker arm, *i.e.*, multifunctional conjugates; at least one peptide being O-glycosylated or including a mutant O-linked glycosylation site. The multi-functional
10 conjugates of the invention can include two or more copies of the same peptide or a collection of diverse peptides with different structures, and/or properties. In exemplary conjugates according to this embodiment, the linker between the two peptides is attached to at least one of the peptides through an O-linked glycosyl residue, such as an O-linked glycosyl intact glycosyl linking group.

15 [0072] The conjugates of the invention are formed by the enzymatic attachment of a modified sugar to the glycosylated or unglycosylated peptide. The modified sugar is directly added to an O-linked glycosylation site, or to a glycosyl residue attached either directly or indirectly (e.g., through one or more glycosyl residue) to an O-linked glycosylation site. The invention also provides a conjugate of an O-linked glycosylated peptide in which a modified
20 sugar is directly attached to an N-linked site, or to a glycosyl residue attached either directly or indirectly to an N-linked glycosylation site.

[0073] The modified sugar, when interposed between the peptide (or glycosyl residue) and the modifying group on the sugar becomes what is referred to herein as “an intact glycosyl linking group.” Using the exquisite selectivity of enzymes, such as glycosyltransferases, the
25 present method provides peptides that bear a desired group at one or more specific locations. Thus, according to the present invention, a modified sugar is attached directly to a selected locus on the peptide chain or, alternatively, the modified sugar is appended onto a carbohydrate moiety of a glycopeptide. Peptides in which modified sugars are bound to both a glycopeptide carbohydrate and directly to an amino acid residue of the peptide backbone
30 are also within the scope of the present invention.

[0074] In contrast to known chemical and enzymatic peptide elaboration strategies, the methods of the invention, make it possible to assemble peptides and glycopeptides that have a

substantially homogeneous derivatization pattern; the enzymes used in the invention are generally selective for a particular amino acid residue or combination of amino acid residues of the peptide. The methods are also practical for large-scale production of modified peptides and glycopeptides. Thus, the methods of the invention provide a practical means for large-scale preparation of glycopeptides having preselected uniform derivatization patterns. The methods are particularly well suited for modification of therapeutic peptides may be used to modify glycopeptides that are incompletely glycosylated during production in cell culture cells (*e.g.*, mammalian cells, insect cells, plant cells, fungal cells, yeast cells, or prokaryotic cells) or transgenic plants or animals. In other embodiments, the invention may be used to glycosylate peptides, such as GLP-1, that are not glycosylated in the wild type state. In still further embodiments, glycosylation sites can be introduced by mutation at any position along the peptide backbone. The invention further provides method for testing the biological activity of mutants with introduced glycosylation sites.

[0075] The methods of the invention also provide conjugates of glycosylated and unglycosylated peptides with increased therapeutic half-life due to, for example, reduced clearance rate, or reduced rate of uptake by the immune or reticuloendothelial system (RES). Moreover, the methods of the invention provide a means for masking antigenic determinants on peptides, thus reducing or eliminating a host immune response against the peptide. Selective attachment of targeting agents to a peptide using an appropriate modified sugar can also be used to target a peptide to a particular tissue or cell surface receptor that is specific for the particular targeting agent. Moreover, there is provided a class of peptides that are specifically modified with a therapeutic moiety conjugated through a glycosyl linking group.

O-Glycosylation

[0076] The present invention provides O-linked glycosylated peptides, conjugates of these species and methods for forming O-linked glycosylated peptides that include a selected amino acid sequence ("an O-linked glycosylation site"). Of particular interest are mutant peptides that include an O-linked glycosylation site that is not present in the wild type peptide. The O-linked glycosylation site is a locus for attachment of a glycosyl residue that bears a modifying group.

[0077] Mucin-type O-linked glycosylation, one of the most abundant forms of protein glycosylation, is found on secreted and cell surface associated glycoproteins of all eukaryotic cells. There is great diversity in the structures created by O-linked glycosylation (hundreds

of potential structures), which are produced by the catalytic activity of hundreds of glycosyltransferase enzymes that are resident in the Golgi complex. Diversity exists at the level of the glycan structure and in positions of attachment of O-glycans to protein backbones. Despite the high degree of potential diversity, it is clear that O-linked glycosylation is a highly regulated process that shows a high degree of conservation among multicellular organisms.

[0078] The first step in mucin-type O-linked glycosylation is catalysed by one or more members of a large family of UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferases (GalNAc-transferases) (EC 2.4.1.41), which transfer GalNAc to serine and threonine acceptor sites (Hassan et al., *J. Biol. Chem.* **275**: 38197-38205 (2000)). To date twelve members of the mammalian GalNAc-transferase family have been identified and characterized (Schwientek et al., *J. Biol. Chem.* **277**: 22623-22638 (2002)), and several additional putative members of this gene family have been predicted from analysis of genome databases. The GalNAc-transferase isoforms have different kinetic properties and show differential expression patterns temporally and spatially, suggesting that they have distinct biological functions (Hassan et al., *J. Biol. Chem.* **275**: 38197-38205 (2000)). Sequence analysis of GalNAc-transferases have led to the hypothesis that these enzymes contain two distinct subunits: a central catalytic unit, and a C-terminal unit with sequence similarity to the plant lectin ricin, designated the "lectin domain" (Hagen et al., *J. Biol. Chem.* **274**: 6797-6803 (1999); Hazes, *Protein Eng.* **10**: 1353-1356 (1997); Breton et al., *Curr. Opin. Struct. Biol.* **9**: 563-571 (1999)). Previous experiments involving site-specific mutagenesis of selected conserved residues confirmed that mutations in the catalytic domain eliminated catalytic activity. In contrast, mutations in the "lectin domain" had no significant effects on catalytic activity of the GalNAc-transferase isoform, GalNAc-T1 (Tenno et al., *J. Biol. Chem.* **277**(49): 47088-96 (2002)). Thus, the C-terminal "lectin domain" was believed not to be functional and not to play roles for the enzymatic functions of GalNAc-transferases (Hagen et al., *J. Biol. Chem.* **274**: 6797-6803 (1999)).

[0079] However, recent evidence demonstrates that some GalNAc-transferases exhibit unique activities with partially GalNAc-glycosylated glycopeptides. The catalytic actions of at least three GalNAc-transferase isoforms, GalNAc-T4, -T7, and -T10, selectively act on glycopeptides corresponding to mucin tandem repeat domains where only some of the clustered potential glycosylation sites have been GalNAc glycosylated by other GalNAc-transferases (Bennett et al., *FEBS Letters* **460**: 226-230 (1999); Ten Hagen et al., *J. Biol.*

Chem. **276**: 17395-17404 (2001); Bennett et al., *J. Biol. Chem.* **273**: 30472-30481 (1998); Ten Hagen et al., *J. Biol. Chem.* **274**: 27867-27874 (1999)). GalNAc-T4 and -T7 recognize different GalNAc-glycosylated peptides and catalyse transfer of GalNAc to acceptor substrate sites in addition to those that were previously utilized. One of the functions of such GalNAc-transferase activities is predicted to represent a control step of the density of O-glycan occupancy in mucins and mucin-like glycoproteins with high density of O-linked glycosylation.

[0080] One example of this is the glycosylation of the cancer-associated mucin MUC1. MUC1 contains a tandem repeat O-linked glycosylated region of 20 residues (HGVTSAPDTRPAPGSTAPPA) with five potential O-linked glycosylation sites. GalNAc-T1, -T2, and -T3 can initiate glycosylation of the MUC1 tandem repeat and incorporate at only three sites (HGVTSAPDTRPAPGSTAPPA, GalNAc attachment sites underlined). GalNAc-T4 is unique in that it is the only GalNAc-transferase isoform identified so far that can complete the O-linked glycan attachment to all five acceptor sites in the 20 amino acid tandem repeat sequence of the breast cancer associated mucin, MUC1. GalNAc-T4 transfers GalNAc to at least two sites not used by other GalNAc-transferase isoforms on the GalNAc₄TAP24 glycopeptide (TAPPAHGVTSAPDTRPAPGSTAPP, unique GalNAc-T4 attachment sites are in bold) (Bennett et al., *J. Biol. Chem.* **273**: 30472-30481 (1998). An activity such as that exhibited by GalNAc-T4 appears to be required for production of the glycoform of MUC1 expressed by cancer cells where all potential sites are glycosylated (Muller et al., *J. Biol. Chem.* **274**: 18165-18172 (1999)). Normal MUC1 from lactating mammary glands has approximately 2.6 O-linked glycans per repeat (Muller et al., *J. Biol. Chem.* **272**: 24780-24793 (1997) and MUC1 derived from the cancer cell line T47D has 4.8 O-linked glycans per repeat (Muller et al., *J. Biol. Chem.* **274**: 18165-18172 (1999)). The cancer-associated form of MUC1 is therefore associated with higher density of O-linked glycan occupancy and this is accomplished by a GalNAc-transferase activity identical to or similar to that of GalNAc-T4.

[0081] Polypeptide GalNAc-transferases, which have not displayed apparent GalNAc-glycopeptide specificities, also appear to be modulated by their putative lectin domains (PCT WO 01/85215 A2). Recently, it was found that mutations in the GalNAc-T1 putative lectin domain, similarly to those previously analysed in GalNAc-T4 (Hassan *et al.*, *J. Biol. Chem.* **275**: 38197-38205 (2000)), modified the activity of the enzyme in a similar fashion as GalNAc-T4. Thus, while wild type GalNAc-T1 added multiple consecutive GalNAc residues

to a peptide substrate with multiple acceptor sites, mutated GalNAc-T1 failed to add more than one GalNAc residue to the same substrate (Tenno *et al.*, *J. Biol. Chem.* 277(49): 47088-96 (2002)).

[0082] Since it has been demonstrated that mutations of GalNAc transferases can be utilized to produce glycosylation patterns that are distinct from those produced by the wild-type enzymes, it is within the scope of the present invention to utilize one or more mutant GalNAc transferase in preparing the O-linked glycosylated peptides of the invention.

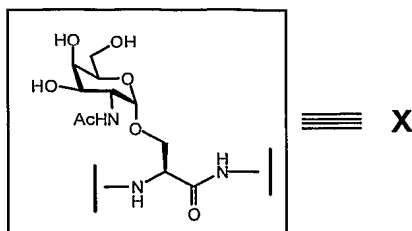
Mutant GLP-1 Peptides with O-linked Glycosylation Sites

[0083] The peptides provided by the present invention include an amino acid sequence that is recognized as a GalNAc acceptor by one or more wild-type or mutant GalNAc transferases. The amino acid sequence of the peptide is either the wild-type, for those peptides that include an O-linked glycosylation site, or may be a mutant sequence in which a non-naturally occurring O-linked glycosylation site is introduced. An exemplary peptide with which the present invention is practiced includes Glucagon-Like Peptide-1 (GLP-1). The emphasis of the following discussion on GLP-1 is for clarity of illustration. Those of skill will understand that the strategy set forth herein for preparing O-linked glycoconjugated analogues of wild-type and mutant peptides is applicable to any peptide.

[0084] In an exemplary embodiment, the peptide is a biologically active GLP-1 mutant that includes one or more mutations at one or more sites distributed along the peptide backbone.

Representative wild type and mutant GLP-1 polypeptides of the invention have sequences that are selected from:

GLP-1 Glycopeptides



Ac-X-HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 H-X-EGTFTSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HA-X-GTFTSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAE-X-TFTSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEG-X-FTSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGT-X-TSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTF-X-SDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFT-X-DVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTS-X-VSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSD-X-SSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDV-X-SYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVS-X-YLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSS-X-LEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSY-X-EGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYL-X-GQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLE-X-QAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEG-X-AAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQ-X-AKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQA-X-KEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAA-X-EFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAK-X-FIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKE-X-IAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEF-X-AWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFI-X-WLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIA-X-LVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAW-X-VKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAWL-X-KGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAWLX-GR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAWLK-X-R-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAWLKX-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR-X-NH₂

[0085] In another exemplary embodiment, the peptide is a fusion of one or more peptides.

In another exemplary embodiment, the components of the peptide are members selected from a GLP-1, GLP-1 analogs and/or GLP-1 mutants. In another exemplary embodiment, the

5 components of the peptide are one or more non-GLP-1 peptides and GLP-1, GLP-1 analog and/or GLP-1 mutant. In another exemplary embodiment, the peptide is an Oxyntomodulin-GLP-1 fusion. This peptide has the following sequence.

HSQGTFTSDY SKYLDSRRAQ DFVQWLMNTK RNRNNIAKRH
DEFERHAEGT FTSDVSSYLE GQAAKEFIW LVKGRG

In another exemplary embodiment, the peptide is an oxyntomodulin/GLP-1 mutant fusion. In another exemplary embodiment, oxyntomodulin/GLP-1 mutant fusions have the following

5 “natural” sequence

-----T²⁹KRNRNNIAKRHDEFERHAE-----, natural sequence;

replaced with sequences that are selected from:

T²⁹**BJJ'**RN(Z')_aNIAOUXX'O'FEZHAE

wherein all substitutions are independently selected from:

10 B = N (natural human variant), K, A, G, S, T, L
J = R, G, A, S, T, L
O = K, P
U = T, S, K
X = H, A, Q, N, G, or any uncharged amino acid
15 X' = D, G, A, N, E, or any uncharged amino acid
Z = R, A, G, S, T, V, I, L or any uncharged amino acid
Z' = G, A
J' = N, S, T
O' = E, A, G, M, any uncharged amino acid
20 a = 0 or 1

Representative examples of oxyntomodulin/GLP-1 mutant fusions have the following natural sequence

-----T²⁹KRNRNNIAKRHDEFERHAE-----, natural sequence;

replaced with sequences that are selected from:

25 --T²⁹NANRRNNIAPTHDEFEAHAE--
--T²⁹NANRRNNIAPTQDEFEAHAE--
--T²⁹NANRRNNIAPTTDEFEAHAE--
--T²⁹NANRRNNIAPTQGEFEAHAE--
--T²⁹NANRRNNIAPTQGAFEAHAE--
30 --T²⁹NANRRNNIAPTQGAMPAHAE--
--T²⁹ARNRRNNIAPTQGAMEAHAE--

--T²⁹NANRRNNIAPTINTFEAHAE--
 --T²⁹NANRRNNIAPTQAYSEGHAE--
 --T²⁹NANRRNNIAPTQAYFEGHAE--
 --T²⁹NANRRNNIAPTTASFEGHAE--
 5 --T²⁹NANRRNNIAPTTLYVEGHAE--
 --T²⁹ARNRRNNIAPTINTFEGHAE--
 --T²⁹NANRRNNIAPTINTFEGHAE--
 --T²⁹NANRSGDIPTINTFEGHAE-

These sequences are based on human sequences in an attempt to minimize immunogenicity while creating a site for glycosylation and preventing proteolysis.

[0086] In another exemplary embodiment, one of the non-GLP-1 peptides is a member selected from a GLP-2, GLP-2 analog and/or GLP-2 mutant. In another exemplary embodiment, the peptide is a GLP-1/GLP-2 fusion. This peptide has the following sequence.

15 HAE GTFTSDVSSY LEGQAAKEFI AWLVKGRGRR DFPEEVAIVE
 ELGRRHADGS FSDEMNTILD NLAARDFINW LIQTKITDRK

In another exemplary embodiment, the peptide is a GLP-1/GLP-2 mutant fusion. In another exemplary embodiment, GLP-1/GLP-2 mutant fusions have the following "natural" sequence

HA--R³⁰GRRDFPEEVAIVEELGRRHADG--, natural sequence;

replaced with sequences that are selected from:

20 HX"--R³⁰GBB'DFPOU(O')_aJJ'VEELGZZ'HADG—

wherein all substitutions are independently selected from:

B and B' (independently selected) = R, A, G, V, I, L, Q, P

J = P, A, I, V, G

O = T, S, E

25 U = E, S, T, Q, I, V, L, and uncharged amino acid

X" = A, G, S, T

Z and Z' (independently selected) = R, A, G, S, T, V, I, L or any uncharged amino acid

J' = E, Y, I, N, A, F, G, or any uncharged amino acid

30 O' = SLP, NT, Y, V, Y

a = 0 or 1

Representative examples of GLP-1/GLP-2 mutant fusions have the following natural sequence

HA--R³⁰GRRDFPEEVAIVEELGRRHADG--, natural sequence;

replaced with sequences that are selected from:

5 HS--R³⁰GQPDFPEGSLPVAIVEELGRGHADG—
 HS--R³⁰GQPDFPTTGSLPVAIVEELGRGHADG—
 HS--R³⁰GQPDFPTTTSEPVAIVEELGRGHADG—
 HS--R³⁰GQPDFPTAVIPVAIVEELGRGHADG—
 HS--R³⁰GQPDFPGSTAPVAIVEELGRGHADG—
 10 HS--R³⁰GQPDFPLTLEPVAIVEELGRGHADG—
 HS--R³⁰GQPDFPTSGEPVAIVEELGRGHADG—
 HS--R³⁰GQPDFPTTINTPVAIVEELGRGHADG—
 HS--R³⁰GQPDFPTTTLYPVAIVEELGRGHADG—
 HS--R³⁰GQPDFPEGSLPTAIVEELGRGHADG—
 15 HS--R³⁰GQPDFPEGSLPTINTEELGRGHADG—
 HS--R³⁰GQPDFPEGSLPTQAVEELGRGHADG--
 HS--R³⁰GQADFPEEVPTIVEELGRGHADG—
 HS--R³⁰GQADFPEEVPTTINTLGRGHADG—
 HS--R³⁰GQADFPEEVPTTQGALGRGHADG—
 20 HS--R³⁰GQADFPEEVPTTTLYLGRGHADG—
 HS--R³⁰GQADFPTTVLPIVEELGRGHADG—
 HS--R³⁰GQADFPTTEIPIVEELGRGHADG—
 HS--R³⁰GQADFPTSDGPIVEELGRGHADG—
 HS--R³⁰GQADFPTTEVPIVEELGRGHADG--

25 These sequences are based on human sequences in an attempt to minimize immunogenicity while creating a site for glycosylation and preventing proteolysis.

[0087] In another exemplary embodiment, the peptide is a fusion of three peptides, thus forming a triple fusion. The three peptides can be arranged in any order. In another exemplary embodiment, the three peptides are oxyntomodulin, GLP-1 and GLP-2. This
 30 peptide has the following sequence.

HSQGTFTS DYSKYLDSRR AQDFVQWLMN TKRNRNNIAK RHDEFERHAE
GTFTSDVSSY LEGQAAKEFI AWLVKGRGRR DFPEEVAIVE ELGRRHADGS
FSDEMNTILD NLAARDFINW LIQTKITDRK

In another exemplary embodiment, the peptide is an oxyntomodulin/GLP-1/GLP-2 mutant fusion. In another exemplary embodiment, an oxyntomodulin/GLP-1/GLP-2 mutant fusion has the following "natural" sequence

5 HS-----T²⁹KRNRNNIAKRHDEFERHAE---H³⁶A--
R⁶⁵GRRDFPEEVAIVEELGRRHADG--, natural sequence;

replaced with sequences that are selected from:

HS--T²⁹**BJJ'**RN(**Z'**)_aNIAOUXX'**O'**FEZHAE--
R⁶⁶**GB'B"**DFPO"**U'** (**O'''**)_a**J''J'''**VEELGX"**Z''**HADG--

wherein all substitutions are independently selected from:

10 B = N (natural human variant), K, A, G, S, T, L
J = R, G, A, S, T, L
O = K, P
U = T, S
X = H, A, Q, N, G, or any uncharged amino acid
15 X' = D, G, A, N, E, or any uncharged amino acid
Z = R, A, G, S, T, V, I, L or any uncharged amino acid
Z' = G, A
J' = N, S, T
O' = E, A, G, M, any uncharged amino acid
20 a = 0 or 1
B'' and B' (independently selected) = R, A, G, V, I, L, Q, P
J'' = P, A, I, V, G
O'' = T, S, E
U' = E, S, T, Q, I, V, L, and uncharged amino acid
25 X'' = A, G, S, T
X''' and Z'' (independently selected) = R, A, G, S, T, V, I, L or any uncharged amino acid
J''' = E, Y, I, N, A, F, G, or any uncharged amino acid
O''' = SLP, NT, Y, V, Y

30 Representative examples of an oxyntomodulin/GLP-1/GLP-2 mutant fusion have the following "natural" sequence

HS-----T²⁹KRNRNNIAKRHDEFERHAE---H³⁶A--
R⁶⁵GRRDFPEEVAIVEELGRRHADG--, natural sequence;

replaced with sequences that are selected from:

HS--T²⁹NANRSGDIPKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQPDFPEGSLPVAIVEELGRGHADG—
 HS--T²⁹NANRSDDIPKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQPDFPEGSLPVAIVEELGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQPDFPTGSLPVAIVEELGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQPDFPTTSEPVAIVEELGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQPDFPTAVIPVAIVEELGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQPDFPGSTAPVAIVEELGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQPDFPLTLEPVAIVEELGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQPDFPTSGEPVAIVEELGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQPDFPTINTPVAIVEELGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQPDFPTTLYPVAIVEELGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQPDFPEGSLPTAIVEELGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQPDFPEGSLPTINTEELGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQPDFPEGSLPTOQAVEELGRGHADG--
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQADFPEEVPTVEELGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQADFPEEVPTIN~~T~~LGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQADFPEEVPTOGALGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQADFPEEVPTTLYLGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQADFPTVLPIVEELGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQADFPTEIPIVEELGRGHADG—
 HS--T²⁹NANANNIAKAHDEF~~E~~AHA~~E~~--
 R⁶⁶GQADFPSDGPIVEELGRGHADG—

HS--T²⁹NANANNIAKAHDEFEAHAE--
R⁶⁶GQADFPTEVPIVEELGRGHADG--

These sequences are based on human sequences in an attempt to minimize immunogenicity while creating a site for glycosylation and preventing proteolysis.

- 5 [0088] Appropriate O-linked glycosylation sequences for GLP-1 and peptides other than GLP-1 can be determined by preparing a polypeptide incorporating a putative O-linked glycosylation site and submitting that polypeptide to suitable O-linked glycosylation conditions, thereby confirming its ability to serve as an acceptor for a GalNAc transferase. Moreover, as will be apparent to one of skill in the art, peptides that include one or more
- 10 mutations are within the scope of the present invention. The mutations are designed to allow the adjustment of desirable properties of the peptides, e.g., activity and number and position of O- and/or N-linked glycosylation sites on the peptide.

Acquisition of Peptide Coding Sequences

General Recombinant Technology

- 15 [0089] This invention relies on routine techniques in the field of recombinant genetics. Basic texts disclosing the general methods of use in this invention include Sambrook and Russell, *Molecular Cloning, A Laboratory Manual* (3rd ed. 2001); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); and Ausubel *et al.*, eds., *Current Protocols in Molecular Biology* (1994).
- 20 [0090] For nucleic acids, sizes are given in either kilobases (kb) or base pairs (bp). These are estimates derived from agarose or acrylamide gel electrophoresis, from sequenced nucleic acids, or from published DNA sequences. For proteins, sizes are given in kilodaltons (kDa) or amino acid residue numbers. Proteins sizes are estimated from gel electrophoresis, from sequenced proteins, from derived amino acid sequences, or from published protein sequences.
- 25 [0091] Oligonucleotides that are not commercially available can be chemically synthesized, e.g., according to the solid phase phosphoramidite triester method first described by Beaucage & Caruthers, *Tetrahedron Lett.* **22**: 1859-1862 (1981), using an automated synthesizer, as described in Van Devanter *et. al.*, *Nucleic Acids Res.* **12**: 6159-6168 (1984). Entire genes can also be chemically synthesized. Purification of oligonucleotides is
- 30 performed using any art-recognized strategy, e.g., native acrylamide gel electrophoresis or anion-exchange HPLC as described in Pearson & Reanier, *J. Chrom.* **255**: 137-149 (1983).

[0092] The sequence of the cloned wild-type peptide genes, polynucleotide encoding mutant peptides, and synthetic oligonucleotides can be verified after cloning using, *e.g.*, the chain termination method for sequencing double-stranded templates of Wallace *et al.*, *Gene* 16: 21-26 (1981).

5 *Cloning and Subcloning of a Wild-Type Peptide Coding Sequence*

[0093] Numerous polynucleotide sequences encoding wild-type peptides have been determined and are available from a commercial supplier, *e.g.*, human growth hormone, *e.g.*, GenBank Accession Nos. NM 000515, NM 002059, NM 022556, NM 022557, NM 022558, NM 022559, NM 022560, NM 022561, and NM 022562.

10 [0094] The rapid progress in the studies of human genome has made possible a cloning approach where a human DNA sequence database can be searched for any gene segment that has a certain percentage of sequence homology to a known nucleotide sequence, such as one encoding a previously identified peptide. Any DNA sequence so identified can be subsequently obtained by chemical synthesis and/or a polymerase chain reaction (PCR)
15 technique such as overlap extension method. For a short sequence, completely *de novo* synthesis may be sufficient; whereas further isolation of full length coding sequence from a human cDNA or genomic library using a synthetic probe may be necessary to obtain a larger gene.

[0095] Alternatively, a nucleic acid sequence encoding a peptide can be isolated from a
20 human cDNA or genomic DNA library using standard cloning techniques such as polymerase chain reaction (PCR), where homology-based primers can often be derived from a known nucleic acid sequence encoding a peptide. Most commonly used techniques for this purpose are described in standard texts, *e.g.*, Sambrook and Russell, *supra*.

[0096] cDNA libraries suitable for obtaining a coding sequence for a wild-type peptide
25 may be commercially available or can be constructed. The general methods of isolating mRNA, making cDNA by reverse transcription, ligating cDNA into a recombinant vector, transfecting into a recombinant host for propagation, screening, and cloning are well known (*see, e.g.*, Gubler and Hoffman, *Gene*, 25: 263-269 (1983); Ausubel *et al.*, *supra*). Upon obtaining an amplified segment of nucleotide sequence by PCR, the segment can be further
30 used as a probe to isolate the full-length polynucleotide sequence encoding the wild-type peptide from the cDNA library. A general description of appropriate procedures can be found in Sambrook and Russell, *supra*.

[0097] A similar procedure can be followed to obtain a full length sequence encoding a wild-type peptide, *e.g.*, any one of the GenBank Accession Nos mentioned above, from a human genomic library. Human genomic libraries are commercially available or can be constructed according to various art-recognized methods. In general, to construct a genomic library, the DNA is first extracted from an tissue where a peptide is likely found. The DNA is then either mechanically sheared or enzymatically digested to yield fragments of about 12-20 kb in length. The fragments are subsequently separated by gradient centrifugation from polynucleotide fragments of undesired sizes and are inserted in bacteriophage λ vectors. These vectors and phages are packaged *in vitro*. Recombinant phages are analyzed by plaque hybridization as described in Benton and Davis, *Science*, **196**: 180-182 (1977). Colony hybridization is carried out as described by Grunstein *et al.*, *Proc. Natl. Acad. Sci. USA*, **72**: 3961-3965 (1975).

[0098] Based on sequence homology, degenerate oligonucleotides can be designed as primer sets and PCR can be performed under suitable conditions (*see, e.g.*, White *et al.*, *PCR Protocols: Current Methods and Applications*, 1993; Griffin and Griffin, *PCR Technology*, CRC Press Inc. 1994) to amplify a segment of nucleotide sequence from a cDNA or genomic library. Using the amplified segment as a probe, the full-length nucleic acid encoding a wild-type peptide is obtained:

[0099] Upon acquiring a nucleic acid sequence encoding a wild-type peptide, the coding sequence can be subcloned into a vector, for instance, an expression vector, so that a recombinant wild-type peptide can be produced from the resulting construct. Further modifications to the wild-type peptide coding sequence, *e.g.*, nucleotide substitutions, may be subsequently made to alter the characteristics of the molecule.

Introducing Mutations into a Peptide Sequence

[0100] From an encoding polynucleotide sequence, the amino acid sequence of a wild-type peptide can be determined. Subsequently, this amino acid sequence may be modified to alter the protein's glycosylation pattern, by introducing additional glycosylation site(s) at various locations in the amino acid sequence.

[0101] Several types of protein glycosylation sites are well known in the art. For instance, in eukaryotes, N-linked glycosylation occurs on the asparagine of the consensus sequence Asn-X_{aa}-Ser/Thr, in which X_{aa} is any amino acid except proline (Kornfeld *et al.*, *Ann Rev Biochem* **54**:631-664 (1985); Kukuruzinska *et al.*, *Proc. Natl. Acad. Sci. USA* **84**:2145-2149

(1987); Herscovics *et al.*, *FASEB J* 7:540-550 (1993); and Orlean, *Saccharomyces* Vol. 3 (1996)). O-linked glycosylation takes place at serine or threonine residues (Tanner *et al.*, *Biochim. Biophys. Acta.* 906:81-91 (1987); and Hounsell *et al.*, *Glycoconj. J.* 13:19-26 (1996)). Other glycosylation patterns are formed by linking glycosylphosphatidylinositol to the carboxyl-terminal carboxyl group of the protein (Takeda *et al.*, *Trends Biochem. Sci.* 20:367-371 (1995); and Udenfriend *et al.*, *Ann. Rev. Biochem.* 64:593-591 (1995). Based on this knowledge, suitable mutations can thus be introduced into a wild-type peptide sequence to form new glycosylation sites.

[0102] Although direct modification of an amino acid residue within a peptide polypeptide sequence may be suitable to introduce a new N-linked or O-linked glycosylation site, more frequently, introduction of a new glycosylation site is accomplished by mutating the polynucleotide sequence encoding a peptide. This can be achieved by using any of known mutagenesis methods, some of which are discussed below. Exemplary modifications to a GLP-1 peptide include those illustrated in SEQ ID NO:____.

[0103] A variety of mutation-generating protocols are established and described in the art. See, e.g., Zhang *et al.*, *Proc. Natl. Acad. Sci. USA*, 94: 4504-4509 (1997); and Stemmer, *Nature*, 370: 389-391 (1994). The procedures can be used separately or in combination to produce variants of a set of nucleic acids, and hence variants of encoded polypeptides. Kits for mutagenesis, library construction, and other diversity-generating methods are commercially available.

[0104] Mutational methods of generating diversity include, for example, site-directed mutagenesis (Botstein and Shortle, *Science*, 229: 1193-1201 (1985)), mutagenesis using uracil-containing templates (Kunkel, *Proc. Natl. Acad. Sci. USA*, 82: 488-492 (1985)), oligonucleotide-directed mutagenesis (Zoller and Smith, *Nucl. Acids Res.*, 10: 6487-6500 (1982)), phosphorothioate-modified DNA mutagenesis (Taylor *et al.*, *Nucl. Acids Res.*, 13: 8749-8764 and 8765-8787 (1985)), and mutagenesis using gapped duplex DNA (Kramer *et al.*, *Nucl. Acids Res.*, 12: 9441-9456 (1984)).

[0105] Other methods for generating mutations include point mismatch repair (Kramer *et al.*, *Cell*, 38: 879-887 (1984)), mutagenesis using repair-deficient host strains (Carter *et al.*, *Nucl. Acids Res.*, 13: 4431-4443 (1985)), deletion mutagenesis (Eghtedarzadeh and Henikoff, *Nucl. Acids Res.*, 14: 5115 (1986)), restriction-selection and restriction-purification (Wells *et al.*, *Phil. Trans. R. Soc. Lond. A*, 317: 415-423 (1986)), mutagenesis by total gene synthesis

(Nambiar *et al.*, *Science*, **223**: 1299-1301 (1984)), double-strand break repair (Mandecki, *Proc. Natl. Acad. Sci. USA*, **83**: 7177-7181 (1986)), mutagenesis by polynucleotide chain termination methods (U.S. Patent No. 5,965,408), and error-prone PCR (Leung *et al.*, *Biotechniques*, **1**: 11-15 (1989)).

5 *Modification of Nucleic Acids for Preferred Codon Usage in a Host Organism*

[0106] The polynucleotide sequence encoding a mutant peptide can be further altered to coincide with the preferred codon usage of a particular host. For example, the preferred codon usage of one strain of bacterial cells can be used to derive a polynucleotide that encodes a mutant peptide of the invention and includes the codons favored by this strain. The
10 frequency of preferred codon usage exhibited by a host cell can be calculated by averaging frequency of preferred codon usage in a large number of genes expressed by the host cell (*e.g.*, calculation service is available from web site of the Kazusa DNA Research Institute, Japan). This analysis is preferably limited to genes that are highly expressed by the host cell. U.S. Patent No. 5,824,864, for example, provides the frequency of codon usage by highly
15 expressed genes exhibited by dicotyledonous plants and monocotyledonous plants.

[0107] At the completion of modification, the mutant peptide coding sequences are verified by sequencing and are then subcloned into an appropriate expression vector for recombinant production in the same manner as the wild-type peptides.

Expression and Purification of the Mutant Peptide

20 [0108] Following sequence verification, the mutant peptide of the present invention can be produced using routine techniques in the field of recombinant genetics, relying on the polynucleotide sequences encoding the polypeptide disclosed herein.

Expression Systems

[0109] To obtain high-level expression of a nucleic acid encoding a mutant peptide of the
25 present invention, one typically subclones a polynucleotide encoding the mutant peptide into an expression vector that contains a strong promoter to direct transcription, a transcription/translation terminator and a ribosome binding site for translational initiation. Suitable bacterial promoters are well known in the art and described, *e.g.*, in Sambrook and Russell, *supra*, and Ausubel *et al.*, *supra*. Bacterial expression systems for expressing the
30 wild-type or mutant peptide are available in, *e.g.*, *E. coli*, *Bacillus sp.*, *Salmonella*, and *Caulobacter*. Kits for such expression systems are commercially available. Eukaryotic expression systems for mammalian cells, yeast, and insect cells are well known in the art and

are also commercially available. In one embodiment, the eukaryotic expression vector is an adenoviral vector, an adeno-associated vector, or a retroviral vector.

[0110] The promoter used to direct expression of a heterologous nucleic acid depends on the particular application. The promoter is optionally positioned about the same distance from the heterologous transcription start site as it is from the transcription start site in its natural setting. As is known in the art, however, some variation in this distance can be accommodated without loss of promoter function.

[0111] In addition to the promoter, the expression vector typically includes a transcription unit or expression cassette that contains all the additional elements required for the expression of the mutant peptide in host cells. A typical expression cassette thus contains a promoter operably linked to the nucleic acid sequence encoding the mutant peptide and signals required for efficient polyadenylation of the transcript, ribosome binding sites, and translation termination. The nucleic acid sequence encoding the peptide is typically linked to a cleavable signal peptide sequence to promote secretion of the peptide by the transformed cell. Such signal peptides include, among others, the signal peptides from tissue plasminogen activator, insulin, and neuron growth factor, and juvenile hormone esterase of *Heliothis virescens*. Additional elements of the cassette may include enhancers and, if genomic DNA is used as the structural gene, introns with functional splice donor and acceptor sites.

[0112] In addition to a promoter sequence, the expression cassette should also contain a transcription termination region downstream of the structural gene to provide for efficient termination. The termination region may be obtained from the same gene as the promoter sequence or may be obtained from different genes.

[0113] The particular expression vector used to transport the genetic information into the cell is not particularly critical. Any of the conventional vectors used for expression in eukaryotic or prokaryotic cells may be used. Standard bacterial expression vectors include plasmids such as pBR322-based plasmids, pSKF, pET23D, and fusion expression systems such as GST and LacZ. Epitope tags can also be added to recombinant proteins to provide convenient methods of isolation, *e.g.*, c-myc.

[0114] Expression vectors containing regulatory elements from eukaryotic viruses are typically used in eukaryotic expression vectors, *e.g.*, SV40 vectors, papilloma virus vectors, and vectors derived from Epstein-Barr virus. Other exemplary eukaryotic vectors include pMSG, pAV009/A⁺, pMTO10/A⁺, pMAMneo-5, baculovirus pDSVE, and any other vector

allowing expression of proteins under the direction of the SV40 early promoter, SV40 later promoter, metallothionein promoter, murine mammary tumor virus promoter, Rous sarcoma virus promoter, polyhedrin promoter, or other promoters shown effective for expression in eukaryotic cells.

- 5 [0115] Some expression systems have markers that provide gene amplification such as thymidine kinase, hygromycin B phosphotransferase, and dihydrofolate reductase. Alternatively, high yield expression systems not involving gene amplification are also suitable, such as a baculovirus vector in insect cells, with a polynucleotide sequence encoding the mutant peptide under the direction of the polyhedrin promoter or other strong baculovirus
10 promoters.

- [0116] The elements that are typically included in expression vectors also include a replicon that functions in *E. coli*, a gene encoding antibiotic resistance to permit selection of bacteria that harbor recombinant plasmids, and unique restriction sites in nonessential regions of the plasmid to allow insertion of eukaryotic sequences. The particular antibiotic resistance
15 gene chosen is not critical, any of the many resistance genes known in the art are suitable. The prokaryotic sequences are optionally chosen such that they do not interfere with the replication of the DNA in eukaryotic cells, if necessary.

- [0117] When periplasmic expression of a recombinant protein (*e.g.*, a high mutant of the present invention) is desired, the expression vector further comprises a sequence encoding a
20 secretion signal, such as the *E. coli* OppA (Periplasmic Oligopeptide Binding Protein) secretion signal or a modified version thereof, which is directly connected to 5' of the coding sequence of the protein to be expressed. This signal sequence directs the recombinant protein produced in cytoplasm through the cell membrane into the periplasmic space. The expression vector may further comprise a coding sequence for signal peptidase 1, which is capable of
25 enzymatically cleaving the signal sequence when the recombinant protein is entering the periplasmic space. More detailed description for periplasmic production of a recombinant protein can be found in, *e.g.*, Gray *et al.*, *Gene* 39: 247-254 (1985), U.S. Patent Nos. 6,160,089 and 6,436,674.

- [0118] As discussed above, a person skilled in the art will recognize that various
30 conservative substitutions can be made to any wild-type or mutant peptide or its coding sequence while still retaining the biological activity of the peptide. Moreover, modifications

of a polynucleotide coding sequence may also be made to accommodate preferred codon usage in a particular expression host without altering the resulting amino acid sequence.

Transfection Methods

[0119] Standard transfection methods are used to produce bacterial, mammalian, yeast or
5 insect cell lines that express large quantities of the mutant peptide, which are then purified using standard techniques (*see, e.g., Colley et al., J. Biol. Chem.* **264**: 17619-17622 (1989); *Guide to Protein Purification*, in *Methods in Enzymology*, vol. 182 (Deutscher, ed., 1990)). Transformation of eukaryotic and prokaryotic cells are performed according to standard techniques (*see, e.g., Morrison, J. Bact.* **132**: 349-351 (1977); Clark-Curtiss & Curtiss,
10 *Methods in Enzymology* **101**: 347-362 (Wu *et al.*, eds, 1983)).

[0120] Any of the well-known procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, polybrene, protoplast fusion, electroporation, liposomes, microinjection, plasma vectors, viral
15 vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA, or other foreign genetic material into a host cell (*see, e.g., Sambrook and Russell, supra*). It is only necessary that the particular genetic engineering procedure used be capable of successfully introducing at least one gene into the host cell capable of expressing the mutant peptide.

Detection of Expression of Mutant Peptide in Host Cells

20 [0121] After the expression vector is introduced into appropriate host cells, the transfected cells are cultured under conditions favoring expression of the mutant peptide. The cells are then screened for the expression of the recombinant polypeptide, which is subsequently recovered from the culture using standard techniques (*see, e.g., Scopes, Protein Purification: Principles and Practice* (1982); U.S. Patent No. 4,673,641; Ausubel *et al., supra*; and
25 Sambrook and Russell, *supra*).

[0122] Several general methods for screening gene expression are well known among those skilled in the art. First, gene expression can be detected at the nucleic acid level. A variety of methods of specific DNA and RNA measurement using nucleic acid hybridization techniques are commonly used (*e.g., Sambrook and Russell, supra*). Some methods involve
30 an electrophoretic separation (*e.g., Southern blot for detecting DNA and Northern blot for detecting RNA*), but detection of DNA or RNA can be carried out without electrophoresis as

well (such as by dot blot). The presence of nucleic acid encoding a mutant peptide in transfected cells can also be detected by PCR or RT-PCR using sequence-specific primers.

[0123] Second, gene expression can be detected at the polypeptide level. Various immunological assays are routinely used by those skilled in the art to measure the level of a gene product, particularly using polyclonal or monoclonal antibodies that react specifically with a mutant peptide of the present invention, such as a polypeptide having the amino acid sequence of SEQ ID NO:1-7, (*e.g.*, Harlow and Lane, *Antibodies, A Laboratory Manual*, Chapter 14, Cold Spring Harbor, 1988; Kohler and Milstein, *Nature*, **256**: 495-497 (1975)). Such techniques require antibody preparation by selecting antibodies with high specificity against the mutant peptide or an antigenic portion thereof. The methods of raising polyclonal and monoclonal antibodies are well established and their descriptions can be found in the literature, *see, e.g.*, Harlow and Lane, *supra*; Kohler and Milstein, *Eur. J. Immunol.*, **6**: 511-519 (1976). More detailed descriptions of preparing antibody against the mutant peptide of the present invention and conducting immunological assays detecting the mutant peptide are provided in a later section.

Purification of Recombinantly Produced Mutant Peptide

[0124] Once the expression of a recombinant mutant peptide in transfected host cells is confirmed, the host cells are then cultured in an appropriate scale for the purpose of purifying the recombinant polypeptide.

1. Purification of Recombinantly Produced Mutant Peptide from Bacteria

[0125] When the mutant peptides of the present invention are produced recombinantly by transformed bacteria in large amounts, typically after promoter induction, although expression can be constitutive, the proteins may form insoluble aggregates. There are several protocols that are suitable for purification of protein inclusion bodies. For example, purification of aggregate proteins (hereinafter referred to as inclusion bodies) typically involves the extraction, separation and/or purification of inclusion bodies by disruption of bacterial cells, *e.g.*, by incubation in a buffer of about 100-150 µg/ml lysozyme and 0.1% Nonidet P40, a non-ionic detergent. The cell suspension can be ground using a Polytron grinder (Brinkman Instruments, Westbury, NY). Alternatively, the cells can be sonicated on ice. Alternate methods of lysing bacteria are described in Ausubel *et al.* and Sambrook and Russell, both *supra*, and will be apparent to those of skill in the art.

[0126] The cell suspension is generally centrifuged and the pellet containing the inclusion bodies resuspended in buffer which does not dissolve but washes the inclusion bodies, *e.g.*, 20 mM Tris-HCl (pH 7.2), 1 mM EDTA, 150 mM NaCl and 2% Triton-X 100, a non-ionic detergent. It may be necessary to repeat the wash step to remove as much cellular debris as possible. The remaining pellet of inclusion bodies may be resuspended in an appropriate buffer (*e.g.*, 20 mM sodium phosphate, pH 6.8, 150 mM NaCl). Other appropriate buffers will be apparent to those of skill in the art.

[0127] Following the washing step, the inclusion bodies are solubilized by the addition of a solvent that is both a strong hydrogen acceptor and a strong hydrogen donor (or a combination of solvents each having one of these properties). The proteins that formed the inclusion bodies may then be renatured by dilution or dialysis with a compatible buffer. Suitable solvents include, but are not limited to, urea (from about 4 M to about 8 M), formamide (at least about 80%, volume/volume basis), and guanidine hydrochloride (from about 4 M to about 8 M). Some solvents that are capable of solubilizing aggregate-forming proteins, such as SDS (sodium dodecyl sulfate) and 70% formic acid, may be inappropriate for use in this procedure due to the possibility of irreversible denaturation of the proteins, accompanied by a lack of immunogenicity and/or activity. Although guanidine hydrochloride and similar agents are denaturants, this denaturation is not irreversible and renaturation may occur upon removal (by dialysis, for example) or dilution of the denaturant, allowing re-formation of the immunologically and/or biologically active protein of interest. After solubilization, the protein can be separated from other bacterial proteins by standard separation techniques. For further description of purifying recombinant peptide from bacterial inclusion body, *see, e.g.*, Patra *et al.*, *Protein Expression and Purification* **18**: 182-190 (2000).

[0128] Alternatively, it is possible to purify recombinant polypeptides, *e.g.*, a mutant peptide, from bacterial periplasm. Where the recombinant protein is exported into the periplasm of the bacteria, the periplasmic fraction of the bacteria can be isolated by cold osmotic shock in addition to other methods known to those of skill in the art (*see e.g.*, Ausubel *et al.*, *supra*). To isolate recombinant proteins from the periplasm, the bacterial cells are centrifuged to form a pellet. The pellet is resuspended in a buffer containing 20% sucrose. To lyse the cells, the bacteria are centrifuged and the pellet is resuspended in ice-cold 5 mM MgSO₄ and kept in an ice bath for approximately 10 minutes. The cell suspension is centrifuged and the supernatant decanted and saved. The recombinant proteins

present in the supernatant can be separated from the host proteins by standard separation techniques well known to those of skill in the art.

2. *Standard Protein Separation Techniques for Purification*

[0129] When a recombinant polypeptide, *e.g.*, the mutant peptide of the present invention, is expressed in host cells in a soluble form, its purification can follow the standard protein purification procedure described below.

i. Solubility Fractionation

[0130] Often as an initial step, and if the protein mixture is complex, an initial salt fractionation can separate many of the unwanted host cell proteins (or proteins derived from the cell culture media) from the recombinant protein of interest, *e.g.*, a mutant peptide of the present invention. The preferred salt is ammonium sulfate. Ammonium sulfate precipitates proteins by effectively reducing the amount of water in the protein mixture. Proteins then precipitate on the basis of their solubility. The more hydrophobic a protein is, the more likely it is to precipitate at lower ammonium sulfate concentrations. A typical protocol is to add saturated ammonium sulfate to a protein solution so that the resultant ammonium sulfate concentration is between 20-30%. This will precipitate the most hydrophobic proteins. The precipitate is discarded (unless the protein of interest is hydrophobic) and ammonium sulfate is added to the supernatant to a concentration known to precipitate the protein of interest. The precipitate is then solubilized in buffer and the excess salt removed if necessary, through either dialysis or diafiltration. Other methods that rely on solubility of proteins, such as cold ethanol precipitation, are well known to those of skill in the art and can be used to fractionate complex protein mixtures.

ii. Size Differential Filtration

[0131] Based on a calculated molecular weight, a protein of greater and lesser size can be isolated using ultrafiltration through membranes of different pore sizes (for example, Amicon or Millipore membranes). As a first step, the protein mixture is ultrafiltered through a membrane with a pore size that has a lower molecular weight cut-off than the molecular weight of a protein of interest, *e.g.*, a mutant peptide. The retentate of the ultrafiltration is then ultrafiltered against a membrane with a molecular cut off greater than the molecular weight of the protein of interest. The recombinant protein will pass through the membrane into the filtrate. The filtrate can then be chromatographed as described below.

iii. Column Chromatography

[0132] The proteins of interest (such as the mutant peptide of the present invention) can also be separated from other proteins on the basis of their size, net surface charge, hydrophobicity, or affinity for ligands. In addition, antibodies raised against peptide can be conjugated to column matrices and the peptide immunopurified. All of these methods are well known in the art.

[0133] It will be apparent to one of skill that chromatographic techniques can be performed at any scale and using equipment from many different manufacturers (*e.g.*, Pharmacia Biotech).

10 Immunoassays for Detection of Mutant Peptide Expression

[0134] To confirm the production of a recombinant mutant peptide, immunological assays may be useful to detect in a sample the expression of the polypeptide. Immunological assays are also useful for quantifying the expression level of the recombinant hormone. Antibodies against a mutant peptide are necessary for carrying out these immunological assays.

15 Production of Antibodies against Mutant Peptide

[0135] Methods for producing polyclonal and monoclonal antibodies that react specifically with an immunogen of interest are known to those of skill in the art (*see, e.g.*, Coligan, *Current Protocols in Immunology* Wiley/Greene, NY, 1991; Harlow and Lane, *Antibodies: A Laboratory Manual* Cold Spring Harbor Press, NY, 1989; Stites *et al.* (eds.) *Basic and Clinical Immunology* (4th ed.) Lange Medical Publications, Los Altos, CA, and references cited therein; Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed.) Academic Press, New York, NY, 1986; and Kohler and Milstein *Nature* **256**: 495-497, 1975). Such techniques include antibody preparation by selection of antibodies from libraries of recombinant antibodies in phage or similar vectors (*see*, Huse *et al.*, *Science* **246**: 1275-1281, 1989; and Ward *et al.*, *Nature* **341**: 544-546, 1989).

[0136] In order to produce antisera containing antibodies with desired specificity, the polypeptide of interest (*e.g.*, a mutant peptide of the present invention) or an antigenic fragment thereof can be used to immunize suitable animals, *e.g.*, mice, rabbits, or primates. A standard adjuvant, such as Freund's adjuvant, can be used in accordance with a standard immunization protocol. Alternatively, a synthetic antigenic peptide derived from that particular polypeptide can be conjugated to a carrier protein and subsequently used as an immunogen.

[0137] The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to the antigen of interest. When appropriately high titers of antibody to the antigen are obtained, blood is collected from the animal and antisera are prepared. Further fractionation of the antisera to enrich antibodies specifically reactive to the antigen and purification of the antibodies can be performed subsequently, *see*, Harlow and Lane, *supra*, and the general descriptions of protein purification provided above.

[0138] Monoclonal antibodies are obtained using various techniques familiar to those of skill in the art. Typically, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (*see*, Kohler and Milstein, *Eur. J. Immunol.* 6:511-519, 1976). Alternative methods of immortalization include, *e.g.*, transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and the yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host.

[0139] Additionally, monoclonal antibodies may also be recombinantly produced upon identification of nucleic acid sequences encoding an antibody with desired specificity or a binding fragment of such antibody by screening a human B cell cDNA library according to the general protocol outlined by Huse *et al.*, *supra*. The general principles and methods of recombinant polypeptide production discussed above are applicable for antibody production by recombinant methods.

[0140] When desired, antibodies capable of specifically recognizing a mutant peptide of the present invention can be tested for their cross-reactivity against the wild-type peptide and thus distinguished from the antibodies against the wild-type protein. For instance, antisera obtained from an animal immunized with a mutant peptide can be run through a column on which a wild-type peptide is immobilized. The portion of the antisera that passes through the column recognizes only the mutant peptide and not the wild-type peptide. Similarly, monoclonal antibodies against a mutant peptide can also be screened for their exclusivity in recognizing only the mutant but not the wild-type peptide.

[0141] Polyclonal or monoclonal antibodies that specifically recognize only the mutant peptide of the present invention but not the wild-type peptide are useful for isolating the

mutant protein from the wild-type protein, for example, by incubating a sample with a mutant peptide-specific polyclonal or monoclonal antibody immobilized on a solid support.

Immunoassays for Detecting Mutant Peptide Expression

[0142] Once antibodies specific for a mutant peptide of the present invention are available, the amount of the polypeptide in a sample, *e.g.*, a cell lysate, can be measured by a variety of immunoassay methods providing qualitative and quantitative results to a skilled artisan. For a review of immunological and immunoassay procedures in general *see, e.g.*, Stites, *supra*; U.S. Patent Nos. 4,366,241; 4,376,110; 4,517,288; and 4,837,168.

Labeling in Immunoassays

[0143] Immunoassays often utilize a labeling agent to specifically bind to and label the binding complex formed by the antibody and the target protein. The labeling agent may itself be one of the moieties comprising the antibody/target protein complex, or may be a third moiety, such as another antibody, that specifically binds to the antibody/target protein complex. A label may be detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Examples include, but are not limited to, magnetic beads (*e.g.*, DynabeadsTM), fluorescent dyes (*e.g.*, fluorescein isothiocyanate, Texas red, rhodamine, and the like), radiolabels (*e.g.*, ³H, ¹²⁵I, ³⁵S, ¹⁴C, or ³²P), enzymes (*e.g.*, horse radish peroxidase, alkaline phosphatase, and others commonly used in an ELISA), and colorimetric labels such as colloidal gold or colored glass or plastic (*e.g.*, polystyrene, polypropylene, latex, *etc.*) beads.

[0144] In some cases, the labeling agent is a second antibody bearing a detectable label. Alternatively, the second antibody may lack a label, but it may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second antibody can be modified with a detectable moiety, such as biotin, to which a third labeled molecule can specifically bind, such as enzyme-labeled streptavidin.

[0145] Other proteins capable of specifically binding immunoglobulin constant regions, such as protein A or protein G, can also be used as the label agents. These proteins are normal constituents of the cell walls of streptococcal bacteria. They exhibit a strong non-immunogenic reactivity with immunoglobulin constant regions from a variety of species (*see, generally*, Kronval, et al. *J. Immunol.*, **111**: 1401-1406 (1973); and Akerstrom, et al., *J. Immunol.*, **135**: 2589-2542 (1985)).

Immunoassay Formats

[0146] Immunoassays for detecting a target protein of interest (*e.g.*, a mutant human growth hormone) from samples may be either competitive or noncompetitive.

Noncompetitive immunoassays are assays in which the amount of captured target protein is directly measured. In one preferred “sandwich” assay, for example, the antibody specific for the target protein can be bound directly to a solid substrate where the antibody is immobilized. It then captures the target protein in test samples. The antibody/target protein complex thus immobilized is then bound by a labeling agent, such as a second or third antibody bearing a label, as described above.

[0147] In competitive assays, the amount of target protein in a sample is measured indirectly by measuring the amount of an added (exogenous) target protein displaced (or competed away) from an antibody specific for the target protein by the target protein present in the sample. In a typical example of such an assay, the antibody is immobilized and the exogenous target protein is labeled. Since the amount of the exogenous target protein bound to the antibody is inversely proportional to the concentration of the target protein present in the sample, the target protein level in the sample can thus be determined based on the amount of exogenous target protein bound to the antibody and thus immobilized.

[0148] In some cases, western blot (immunoblot) analysis is used to detect and quantify the presence of a mutant peptide in the samples. The technique generally comprises separating sample proteins by gel electrophoresis on the basis of molecular weight, transferring the separated proteins to a suitable solid support (such as a nitrocellulose filter, a nylon filter, or a derivatized nylon filter) and incubating the samples with the antibodies that specifically bind the target protein. These antibodies may be directly labeled or alternatively may be subsequently detected using labeled antibodies (*e.g.*, labeled sheep anti-mouse antibodies) that specifically bind to the antibodies against a mutant peptide.

[0149] Other assay formats include liposome immunoassays (LIA), which use liposomes designed to bind specific molecules (*e.g.*, antibodies) and release encapsulated reagents or markers. The released chemicals are then detected according to standard techniques (*see*, Monroe *et al.*, *Amer. Clin. Prod. Rev.*, **5**: 34-41 (1986)).

The Conjugates

[0150] In a representative aspect, the present invention provides a glycoconjugate between a peptide and a selected modifying group, in which the modifying group is conjugated to the peptide through a glycosyl linking group, *e.g.*, an intact glycosyl linking group. The glycosyl

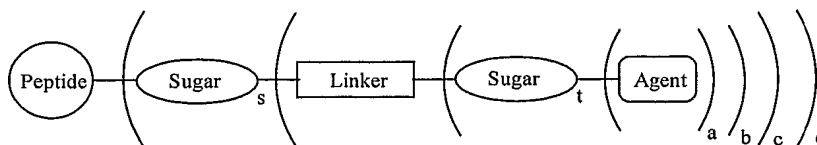
linking group is directly bound to an O-linked glycosylation site on the peptide or, alternatively, it is bound to an O-linked glycosylation site through one or more additional glycosyl residues. Methods of preparing the conjugates are set forth herein and in U.S. Patent No. 5,876,980; 6,030,815; 5,728,554; 5,922,577; WO 98/31826; US2003180835; and WO 03/031464.

[0151] Exemplary peptides include an O-linked GalNAc residue that is bound to the O-linked glycosylation site through the action of a GalNAc transferase. The GalNAc itself may be the intact glycosyl linking group. The GalNAc may also be further elaborated by, for example, a Gal or Sia residue, either of which can act as the intact glycosyl linking group. In representative embodiments, the O-linked saccharyl residue is GalNAc-X, GalNAc-Gal-Sia-X, or GalNAc-Gal-Gal-Sia-X, in which X is a modifying group.

[0152] In an exemplary embodiment, the peptide is a mutant peptide that includes an O-linked glycosylation site not present in the wild-type peptide. The peptide is preferably O-glycosylated at the mutated site with a GalNAc residue. The discussion immediately preceding regarding the structure of the saccharyl moiety is relevant here as well.

[0153] The link between the peptide and the selected moiety includes an intact glycosyl linking group interposed between the peptide and the modifying moiety. As discussed herein, the selected moiety is essentially any species that can be attached to a saccharide unit, resulting in a "modified sugar" that is recognized by an appropriate transferase enzyme, which appends the modified sugar onto the peptide. The saccharide component of the modified sugar, when interposed between the peptide and a selected moiety, becomes an "intact glycosyl linking group." The glycosyl linking group is formed from any mono- or oligo-saccharide that, after modification with a selected moiety, is a substrate for an appropriate transferase.

[0154] The conjugates of the invention will typically correspond to the general structure:



in which the symbols a, b, c, d and s represent a positive, non-zero integer; and t is either 0 or a positive integer. The "agent" is a therapeutic agent, a bioactive agent, a detectable label, water-soluble moiety or the like. The "agent" can be a peptide, e.g., enzyme, antibody, antigen, etc. The linker can be any of a wide array of linking groups, *infra*. Alternatively,

the linker may be a single bond or a "zero order linker." The identity of the peptide is without limitation.

[0155] In an exemplary embodiment, the selected moiety is a water-soluble polymer, e.g., PEG, m-PEG, PPG, m-PPG, etc. The water-soluble polymer is covalently attached to the peptide via a glycosyl linking group. The glycosyl linking group is covalently attached to either an amino acid residue or a glycosyl residue of the peptide. Alternatively, the glycosyl linking group is attached to one or more glycosyl units of a glycopeptide. The invention also provides conjugates in which the glycosyl linking group (e.g., GalNAc) is attached to an amino acid residue (e.g., Thr or Ser).

[0156] In an exemplary embodiment, the protein is an interferon. The interferons are antiviral glycoproteins that, in humans, are secreted by human primary fibroblasts after induction with virus or double-stranded RNA. Interferons are of interest as therapeutics, e.g., antiviral agents (e.g., hepatitis B and C), antitumor agents (e.g., hepatocellular carcinoma) and in the treatment of multiple sclerosis. For references relevant to interferon- α , see, Asano, *et al.*, *Eur. J. Cancer*, **27(Suppl 4)**:S21-S25 (1991); Nagy, *et al.*, *Anticancer Research*, **8(3)**:467-470 (1988); Dron, *et al.*, *J. Biol. Regul. Homeost. Agents*, **3(1)**:13-19 (1989); Habib, *et al.*, *Am. Surg.*, **67(3)**:257-260 (3/2001); and Sugiyama, *et al.*, *Eur. J. Biochem.*, **217**:921-927 (1993). For references discussing interfereon- β , see, e.g., Yu, *et al.*, *J. Neuroimmunol.*, **64(1)**:91-100 (1996); Schmidt, J., *J. Neurosci. Res.*, **65(1)**:59-67 (2001); Wender, *et al.*, *Folia Neuropathol.*, **39(2)**:91-93 (2001); Martin, *et al.*, *Springer Semin. Immunopathol.*, **18(1)**:1-24 (1996); Takane, *et al.*, *J. Pharmacol. Exp. Ther.*, **294(2)**:746-752 (2000); Sburlati, *et al.*, *Biotechnol. Prog.*, **14**:189-192 (1998); Dodd, *et al.*, *Biochimica et Biophysica Acta*, **787**:183-187 (1984); Edelbaum, *et al.*, *J. Interferon Res.*, **12**:449-453 (1992); Conradt, *et al.*, *J. Biol. Chem.*, **262(30)**:14600-14605 (1987); Civas, *et al.*, *Eur. J. Biochem.*, **173**:311-316 (1988); Demolder, *et al.*, *J. Biotechnol.*, **32**:179-189 (1994); Sedmak, *et al.*, *J. Interferon Res.*, **9(Suppl 1)**:S61-S65 (1989); Kagawa, *et al.*, *J. Biol. Chem.*, **263(33)**:17508-17515 (1988); Hershenson, *et al.*, U.S. Patent No. 4,894,330; Jayaram, *et al.*, *J. Interferon Res.*, **3(2)**:177-180 (1983); Menge, *et al.*, *Develop. Biol. Standard.*, **66**:391-401 (1987); Vonk, *et al.*, *J. Interferon Res.*, **3(2)**:169-175 (1983); and Adolf, *et al.*, *J. Interferon Res.*, **10**:255-267 (1990).

[0157] In an exemplary interferon conjugate, interferon alpha, e.g., interferon alpha 2 β , is conjugated to a water soluble polymer through an intact glycosyl linker.

[0158] In a further exemplary embodiment, the invention provides a conjugate of human Glucagon-like peptide-1 (GLP-1). GLP-1 is protein that has pleiotropic effects in the maintenance of glyceic control of the organism. GLP-1 is released in response to the oral ingestion of food. GLP-1 appears to regulate plasma glucose levels by a variety of mechanisms including the enhancement of glucose dependent insulin secretion, stimulation of proinsulin gene expression, suppression of glucagon release and gastric emptying, enhancement of insulin sensitivity, and increase of satiety (*see e.g.*, Xiao *et al.* (2001 Biochemisrty 40:2860, and Perfetti, R. and Merkel, P. (2000) European J. of Endocrinology 143:717). GLP-1 is is rapidly cleared from the body. The half life of GLP-1 in vivo is about 5 minutes, with clearance completed within about 12-13 minutes. Even when administered subcutaneously, GLP-1 is cleared from the circulation within 90 minutes (Perfetti, R. and Merkel, P. *supra*)

[0159] In addition to providing conjugates that are formed through an enzymatically added intact glycosyl linking group, the present invention provides conjugates that are highly homogenous in their substitution patterns. Using the methods of the invention, it is possible to form peptide conjugates in which essentially all of the modified sugar moieties across a population of conjugates of the invention are attached to a structurally identical amino acid or glycosyl residue. Thus, in a second aspect, the invention provides a peptide conjugate having a population of water-soluble polymer moieties, which are covalently bound to the peptide through an intact glycosyl linking group. In a preferred conjugate of the invention, essentially each member of the population is bound via the glycosyl linking group to a glycosyl residue of the peptide, and each glycosyl residue of the peptide to which the glycosyl linking group is attached has the same structure.

[0160] Also provided is a peptide conjugate having a population of water-soluble polymer moieties covalently bound thereto through a glycosyl linking group. In a preferred embodiment, essentially every member of the population of water soluble polymer moieties is bound to an amino acid residue of the peptide via an intact glycosyl linking group, and each amino acid residue having an intact glycosyl linking group attached thereto has the same structure.

[0161] The present invention also provides conjugates analogous to those described above in which the peptide is conjugated to a therapeutic moiety, diagnostic moiety, targeting moiety, toxin moiety or the like via a glycosyl linking group. Each of the above-recited moieties can be a small molecule, natural polymer (e.g., polypeptide) or synthetic polymer.

[0162] The conjugates of the invention can include glycosyl linking groups that are mono- or multi-valent (e.g., antennary structures). Thus, conjugates of the invention include both species in which a selected moiety is attached to a peptide via a monovalent glycosyl linking group. Also included within the invention are conjugates in which more than one selected moiety is attached to a peptide via a multivalent linking group.

The Methods

[0163] In addition to the conjugates discussed above, the present invention provides methods for preparing these and other conjugates. Moreover, the invention provides methods of preventing, curing or ameliorating a disease state by administering a conjugate of the invention to a subject at risk of developing a disease or condition, (e.g., diabetes or obesity) or a subject that has the disease or condition.

[0164] Thus, the invention provides a method of forming a covalent conjugate between a selected moiety and a peptide. In exemplary embodiments, the conjugate is formed between a water-soluble polymer, a therapeutic moiety, targeting moiety or a biomolecule, and a glycosylated or non-glycosylated peptide. The polymer, therapeutic moiety or biomolecule is conjugated to the peptide via a glycosyl linking group, which is interposed between, and covalently linked to both the peptide and the modifying group (e.g. water-soluble polymer). The method includes contacting the peptide with a mixture containing a modified sugar and a glycosyltransferase for which the modified sugar is a substrate. The reaction is conducted under conditions appropriate to form a covalent bond between the modified sugar and the peptide. The sugar moiety of the modified sugar is preferably selected from nucleotide sugars, activated sugars and sugars, which are neither nucleotides nor activated.

[0165] The acceptor peptide (O-glycosylated or non-glycosylated) is typically synthesized *de novo*, or recombinantly expressed in a prokaryotic cell (e.g., bacterial cell, such as *E. coli*) or in a eukaryotic cell such as a mammalian, yeast, insect, fungal or plant cell. The peptide can be either a full-length protein or a fragment. Moreover, the peptide can be a wild type or mutated peptide. In an exemplary embodiment, the peptide includes a mutation that adds one or more N- or O-linked glycosylation sites to the peptide sequence.

[0166] The method of the invention also provides for modification of incompletely glycosylated peptides that are produced recombinantly. Many recombinantly produced glycoproteins are incompletely glycosylated, exposing carbohydrate residues that may have undesirable properties, e.g., immunogenicity, recognition by the RES. Employing a modified

sugar in a method of the invention, the peptide can be simultaneously further glycosylated and derivatized with, e.g., a water-soluble polymer, therapeutic agent, or the like. The sugar moiety of the modified sugar can be the residue that would properly be conjugated to the acceptor in a fully glycosylated peptide, or another sugar moiety with desirable properties.

5 [0167] Any peptides modified by the methods of the invention. However, the peptides are typically mutated peptides, produced by methods known in the art, such as site-directed mutagenesis. Glycosylation of peptides is typically either N-linked or O-linked. An exemplary N-linkage is the attachment of the modified sugar to the side chain of an asparagine residue. The tripeptide sequences asparagine-X-serine and asparagine-X-
10 threonine, where X is any amino acid except proline, are the recognition sequences for enzymatic attachment of a carbohydrate moiety to the asparagine side chain. Thus, the presence of either of these tripeptide sequences in a polypeptide creates a potential glycosylation site. O-linked glycosylation refers to the attachment of one sugar (*e.g.*, N-acetylgalactosamine, galactose, mannose, GlcNAc, glucose, fucose or xylose) to the hydroxy
15 side chain of a hydroxyamino acid, preferably serine or threonine, although unusual or non-natural amino acids, *e.g.*, 5-hydroxyproline or 5-hydroxylysine may also be used.

[0168] Moreover, in addition to peptides, the methods of the present invention can be practiced with other biological structures (*e.g.*, glycolipids, lipids, sphingoids, ceramides, whole cells, and the like, containing an O-linked glycosylation site).

20 [0169] Addition of glycosylation sites to a peptide or other structure is conveniently accomplished by altering the amino acid sequence such that it contains one or more glycosylation sites. The addition may also be made by the incorporation of one or more species presenting an –OH group, preferably serine or threonine residues, within the sequence of the peptide (for O-linked glycosylation sites). The addition may be made by mutation or
25 by full chemical synthesis of the peptide. The peptide amino acid sequence is preferably altered through changes at the DNA level, particularly by mutating the DNA encoding the peptide at preselected bases such that codons are generated that will translate into the desired amino acids. The DNA mutation(s) are preferably made using methods known in the art.

[0170] In an exemplary embodiment, the glycosylation site is added by shuffling
30 polynucleotides. Polynucleotides encoding a candidate peptide can be modulated with DNA shuffling protocols. DNA shuffling is a process of recursive recombination and mutation, performed by random fragmentation of a pool of related genes, followed by reassembly of the

fragments by a polymerase chain reaction-like process. *See, e.g.,* Stemmer, *Proc. Natl. Acad. Sci. USA* 91:10747-10751 (1994); Stemmer, *Nature* 370:389-391 (1994); and U.S. Patent Nos. 5,605,793, 5,837,458, 5,830,721 and 5,811,238.

[0171] The present invention also provides means of adding (or removing) one or more
5 selected glycosyl residues to a peptide, after which a modified sugar is conjugated to at least one of the selected glycosyl residues of the peptide. The present embodiment is useful, for example, when it is desired to conjugate the modified sugar to a selected glycosyl residue that is either not present on a peptide or is not present in a desired amount. Thus, prior to coupling a modified sugar to a peptide, the selected glycosyl residue is conjugated to the
10 peptide by enzymatic or chemical coupling. In another embodiment, the glycosylation pattern of a glycopeptide is altered prior to the conjugation of the modified sugar by the removal of a carbohydrate residue from the glycopeptide. *See, for example* WO 98/31826.

[0172] Addition or removal of any carbohydrate moieties present on the glycopeptide is accomplished either chemically or enzymatically. Chemical deglycosylation is preferably
15 brought about by exposure of the polypeptide variant to the compound trifluoromethanesulfonic acid, or an equivalent compound. This treatment results in the cleavage of most or all sugars except the linking sugar (N-acetylglucosamine or N-acetylgalactosamine), while leaving the peptide intact. Chemical deglycosylation is described by Hakimuddin *et al.*, *Arch. Biochem. Biophys.* **259**: 52 (1987) and by Edge *et al.*,
20 *Anal. Biochem.* **118**: 131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptide variants can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura *et al.*, *Meth. Enzymol.* **138**: 350 (1987).

[0173] Chemical addition of glycosyl moieties is carried out by any art-recognized method. Enzymatic addition of sugar moieties is preferably achieved using a modification of the
25 methods set forth herein, substituting native glycosyl units for the modified sugars used in the invention. Other methods of adding sugar moieties are disclosed in U.S. Patent No. 5,876,980, 6,030,815, 5,728,554, and 5,922,577.

[0174] Exemplary attachment points for selected glycosyl residue include, but are not limited to: (a) consensus sites for N-linked glycosylation, and sites for O-linked
30 glycosylation; (b) terminal glycosyl moieties that are acceptors for a glycosyltransferase; (c) arginine, asparagine and histidine; (d) free carboxyl groups; (e) free sulfhydryl groups such as those of cysteine; (f) free hydroxyl groups such as those of serine, threonine, or

hydroxyproline; (g) aromatic residues such as those of phenylalanine, tyrosine, or tryptophan; or (h) the amide group of glutamine. Exemplary methods of use in the present invention are described in WO 87/05330 published Sep. 11, 1987, and in Aplin and Wriston, CRC CRIT. REV. BIOCHEM., pp. 259-306 (1981).

5 [0175] In one embodiment, the invention provides a method for linking two or more peptides through a linking group. The linking group is of any useful structure and may be selected from straight- and branched-chain structures. Preferably, each terminus of the linker, which is attached to a peptide, includes a modified sugar (i.e., a nascent intact glycosyl linking group).

10 [0176] In an exemplary method of the invention, two peptides are linked together via a linker moiety that includes a PEG linker. The construct conforms to the general structure set forth in the cartoon above. As described herein, the construct of the invention includes two intact glycosyl linking groups (i.e., $s + t = 1$). The focus on a PEG linker that includes two glycosyl groups is for purposes of clarity and should not be interpreted as limiting the identity
15 of linker arms of use in this embodiment of the invention.

[0177] Thus, a PEG moiety is functionalized at a first terminus with a first glycosyl unit and at a second terminus with a second glycosyl unit. The first and second glycosyl units are preferably substrates for different transferases, allowing orthogonal attachment of the first and second peptides to the first and second glycosyl units, respectively. In practice, the
20 (glycosyl)¹-PEG-(glycosyl)² linker is contacted with the first peptide and a first transferase for which the first glycosyl unit is a substrate, thereby forming (peptide)¹-(glycosyl)¹-PEG-(glycosyl)². Transferase and/or unreacted peptide is then optionally removed from the reaction mixture. The second peptide and a second transferase for which the second glycosyl unit is a substrate are added to the
25 (peptide)¹-(glycosyl)¹-PEG-(glycosyl)² conjugate, forming (peptide)¹-(glycosyl)¹-PEG-(glycosyl)²-(peptide)²; at least one of the glycosyl residues is either directly or indirectly O-linked. Those of skill in the art will appreciate that the method outlined above is also applicable to forming conjugates between more than two peptides by, for example, the use of a branched PEG, dendrimer, poly(amino acid), polysaccharide or the
30 like.

[0178] The use of reactive derivatives of PEG (or other linkers) to attach one or more peptide moieties to a linker is within the scope of the present invention. The invention is not

limited by the identity of the reactive PEG analogue. Many activated derivatives of poly(ethyleneglycol) are available commercially and in the literature. It is well within the abilities of one of skill to choose, and synthesize if necessary, an appropriate activated PEG derivative with which to prepare a substrate useful in the present invention. *See*, Abuchowski

5 *et al. Cancer Biochem. Biophys.*, **7**: 175-186 (1984); Abuchowski *et al.*, *J. Biol. Chem.*, **252**: 3582-3586 (1977); Jackson *et al.*, *Anal. Biochem.*, **165**: 114-127 (1987); Koide *et al.*, *Biochem Biophys. Res. Commun.*, **111**: 659-667 (1983)), tresylate (Nilsson *et al.*, *Methods Enzymol.*, **104**: 56-69 (1984); Delgado *et al.*, *Biotechnol. Appl. Biochem.*, **12**: 119-128 (1990)); N-hydroxysuccinimide derived active esters (Buckmann *et al.*, *Makromol. Chem.*,

10 **182**: 1379-1384 (1981); Joppich *et al.*, *Makromol. Chem.*, **180**: 1381-1384 (1979); Abuchowski *et al.*, *Cancer Biochem. Biophys.*, **7**: 175-186 (1984); Katre *et al. Proc. Natl. Acad. Sci. U.S.A.*, **84**: 1487-1491 (1987); Kitamura *et al.*, *Cancer Res.*, **51**: 4310-4315 (1991); Boccu *et al.*, *Z. Naturforsch.*, **38C**: 94-99 (1983), carbonates (Zalipsky *et al.*, *POLY(ETHYLENE GLYCOL) CHEMISTRY: BIOTECHNICAL AND BIOMEDICAL APPLICATIONS*, Harris, Ed., Plenum Press, New York, 1992, pp. 347-370; Zalipsky *et al.*, *Biotechnol. Appl. Biochem.*, **15**: 100-114 (1992); Veronese *et al.*, *Appl. Biochem. Biotech.*, **11**: 141-152 (1985)), imidazolyl formates (Beauchamp *et al.*, *Anal. Biochem.*, **131**: 25-33 (1983); Berger *et al.*, *Blood*, **71**: 1641-1647 (1988)), 4-dithiopyridines (Woghiren *et al.*, *Bioconjugate Chem.*, **4**: 314-318 (1993)), isocyanates (Byun *et al.*, *ASAIJ Journal*, M649-M-653 (1992))

20 and epoxides (U.S. Pat. No. 4,806,595, issued to Noishiki *et al.*, (1989). Other linking groups include the urethane linkage between amino groups and activated PEG. *See*, Veronese, *et al.*, *Appl. Biochem. Biotechnol.*, **11**: 141-152 (1985).

[0179] In another exemplary embodiment in which a reactive PEG derivative is utilized, the invention provides a method for extending the blood-circulation half-life of a selected

25 peptide, in essence targeting the peptide to the blood pool, by conjugating the peptide to a synthetic or natural polymer of a size sufficient to retard the filtration of the protein by the glomerulus (e.g., albumin). For example, GLP-1 can be conjugated to albumin via a PEG linker using a combination of chemical and enzymatic modification.

Modified Sugars

30 [0180] Modified glycosyl donor species ("modified sugars") are preferably selected from modified sugar nucleotides, activated modified sugars and modified sugars that are simple saccharides that are neither nucleotides nor activated. Any desired carbohydrate structure can

be added to a peptide using the methods of the invention. Typically, the structure will be a monosaccharide, but the present invention is not limited to the use of modified monosaccharide sugars; oligosaccharides and polysaccharides are useful as well.

[0181] The modifying group is attached to a sugar moiety by enzymatic means, chemical means or a combination thereof, thereby producing a modified sugar. The sugars are substituted at any position that allows for the attachment of the modifying moiety, yet which still allows the sugar to function as a substrate for the enzyme used to ligate the modified sugar to the peptide. In a preferred embodiment, when sialic acid is the sugar, the sialic acid is substituted with the modifying group at either the 9-position on the pyruvyl side chain or at the 5-position on the amine moiety that is normally acetylated in sialic acid.

[0182] In certain embodiments of the present invention, a modified sugar nucleotide is utilized to add the modified sugar to the peptide. Exemplary sugar nucleotides that are used in the present invention in their modified form include nucleotide mono-, di- or triphosphates or analogs thereof. In a preferred embodiment, the modified sugar nucleotide is selected from a UDP-glycoside, CMP-glycoside, or a GDP-glycoside. Even more preferably, the modified sugar nucleotide is selected from an UDP-galactose, UDP-galactosamine, UDP-glucose, UDP-glucosamine, GDP-mannose, GDP-fucose, CMP-sialic acid, or CMP-NeuAc. N-acetylamine derivatives of the sugar nucleotides are also of use in the method of the invention.

[0183] The invention also provides methods for synthesizing a modified peptide using a modified sugar, *e.g.*, modified-galactose, -fucose, -GalNAc and -sialic acid. When a modified sialic acid is used, either a sialyltransferase or a trans-sialidase (for α 2,3-linked sialic acid only) can be used in these methods.

[0184] In other embodiments, the modified sugar is an activated sugar. Activated modified sugars, which are useful in the present invention are typically glycosides which have been synthetically altered to include an activated leaving group. As used herein, the term "activated leaving group" refers to those moieties, which are easily displaced in enzyme-regulated nucleophilic substitution reactions. Many activated sugars are known in the art. See, for example, Vocadlo *et al.*, In CARBOHYDRATE CHEMISTRY AND BIOLOGY, Vol. 2, Ernst *et al.* Ed., Wiley-VCH Verlag: Weinheim, Germany, 2000; Kodama *et al.*, *Tetrahedron Lett.* **34**: 6419 (1993); Loughheed, *et al.*, *J. Biol. Chem.* **274**: 37717 (1999)).

[0185] Examples of activating groups (leaving groups) include fluoro, chloro, bromo, tosylate ester, mesylate ester, triflate ester and the like. Preferred activated leaving groups, for use in the present invention, are those that do not significantly sterically encumber the enzymatic transfer of the glycoside to the acceptor. Accordingly, preferred embodiments of activated glycoside derivatives include glycosyl fluorides and glycosyl mesylates, with glycosyl fluorides being particularly preferred. Among the glycosyl fluorides, α -galactosyl fluoride, α -mannosyl fluoride, α -glucosyl fluoride, α -fucosyl fluoride, α -xylosyl fluoride, α -sialyl fluoride, α -N-acetylglucosaminyl fluoride, α -N-acetylgalactosaminyl fluoride, β -galactosyl fluoride, β -mannosyl fluoride, β -glucosyl fluoride, β -fucosyl fluoride, β -xylosyl fluoride, β -sialyl fluoride, β -N-acetylglucosaminyl fluoride and β -N-acetylgalactosaminyl fluoride are most preferred.

[0186] By way of illustration, glycosyl fluorides can be prepared from the free sugar by first acetylating the sugar and then treating it with HF/pyridine. This generates the thermodynamically most stable anomer of the protected (acetylated) glycosyl fluoride (*i.e.*, the α -glycosyl fluoride). If the less stable anomer (*i.e.*, the β -glycosyl fluoride) is desired, it can be prepared by converting the peracetylated sugar with HBr/HOAc or with HCl to generate the anomeric bromide or chloride. This intermediate is reacted with a fluoride salt such as silver fluoride to generate the glycosyl fluoride. Acetylated glycosyl fluorides may be deprotected by reaction with mild (catalytic) base in methanol (*e.g.* NaOMe/MeOH). In addition, many glycosyl fluorides are commercially available.

[0187] Other activated glycosyl derivatives can be prepared using conventional methods known to those of skill in the art. For example, glycosyl mesylates can be prepared by treatment of the fully benzylated hemiacetal form of the sugar with mesyl chloride, followed by catalytic hydrogenation to remove the benzyl groups.

[0188] In a further exemplary embodiment, the modified sugar is an oligosaccharide having an antennary structure. In a preferred embodiment, one or more of the termini of the antennae bear the modifying moiety. When more than one modifying moiety is attached to an oligosaccharide having an antennary structure, the oligosaccharide is useful to "amplify" the modifying moiety; each oligosaccharide unit conjugated to the peptide attaches multiple copies of the modifying group to the peptide. The general structure of a typical conjugate of the invention as set forth in the drawing above, encompasses multivalent species resulting from preparing a conjugate of the invention utilizing an antennary structure. Many antennary

saccharide structures are known in the art, and the present method can be practiced with them without limitation.

[0189] Exemplary modifying groups are discussed below. The modifying groups can be selected for their ability to impart to a peptide one or more desirable property. Exemplary properties include, but are not limited to, enhanced pharmacokinetics, enhanced pharmacodynamics, improved biodistribution, providing a polyvalent species, improved water solubility, enhanced or diminished lipophilicity, and tissue targeting.

Water-Soluble Polymers

[0190] The hydrophilicity of a selected peptide is enhanced by conjugation with polar molecules such as amine-, ester-, hydroxyl- and polyhydroxyl-containing molecules. Representative examples include, but are not limited to, polylysine, polyethyleneimine, and polyethers, e.g., poly(ethyleneglycol), m-poly(ethylene glycol), poly(propyleneglycol), m-poly(ethylene glycol), and other O-alkyl poly(alkylene glycol) moieties. Preferred water-soluble polymers are essentially non-fluorescent, or emit such a minimal amount of fluorescence that they are inappropriate for use as a fluorescent marker in an assay. Moreover, it is generally preferred to use polymers that are not naturally occurring sugars. An exception to this preference is the use of an otherwise naturally occurring sugar that is modified by covalent attachment of another entity (e.g., poly(ethylene glycol), poly(propylene glycol), biomolecule, therapeutic moiety, diagnostic moiety, etc.). In another exemplary embodiment, a therapeutic sugar moiety is conjugated to a linker arm and the sugar-linker arm cassette is subsequently conjugated to a peptide via a method of the invention.

[0191] Methods and chemistry for activation of water-soluble polymers and saccharides as well as methods for conjugating saccharides and polymers to various species are described in the literature. Commonly used methods for activation of polymers include activation of functional groups with cyanogen bromide, periodate, glutaraldehyde, biepoxydes, epichlorohydrin, divinylsulfone, carbodiimide, sulfonyl halides, trichlorotriazine, etc. (see, R. F. Taylor, (1991), PROTEIN IMMOBILISATION. FUNDAMENTALS AND APPLICATIONS, Marcel Dekker, N.Y.; S. S. Wong, (1992), CHEMISTRY OF PROTEIN CONJUGATION AND CROSSLINKING, CRC Press, Boca Raton; G. T. Hermanson *et al.*, (1993), IMMOBILIZED AFFINITY LIGAND TECHNIQUES, Academic Press, N.Y.; Dunn, R.L., *et al.*, Eds. POLYMERIC

DRUGS AND DRUG DELIVERY SYSTEMS, ACS Symposium Series Vol. 469, American Chemical Society, Washington, D.C. 1991).

[0192] Many water-soluble polymers are known to those of skill in the art and are useful in practicing the present invention. The term water-soluble polymer encompasses species such as saccharides (e.g., dextran, amylose, hyaluronic acid, poly(sialic acid), heparans, heparins, etc.); poly (amino acids); nucleic acids; synthetic polymers (e.g., poly(acrylic acid), poly(ethers), e.g., poly(ethylene glycol); peptides, proteins, and the like. The present invention may be practiced with any water-soluble polymer with the sole limitation that the polymer must include a point at which the remainder of the conjugate can be attached.

[0193] Methods for activation of polymers can also be found in WO 94/17039, U.S. Pat. No. 5,324,844, WO 94/18247, WO 94/04193, U.S. Pat. No. 5,219,564, U.S. Pat. No. 5,122,614, WO 90/13540, U.S. Pat. No. 5,281,698, and more WO 93/15189, and for conjugation between activated polymers and peptides, e.g. Coagulation Factor VIII (WO 94/15625), haemoglobin (WO 94/09027), oxygen carrying molecule (U.S. Pat. No. 4,412,989), ribonuclease and superoxide dismutase (Veronese *et al.*, *App. Biochem. Biotech.* **11**: 141-45 (1985)).

[0194] Preferred water-soluble polymers are those in which a substantial proportion of the polymer molecules in a sample of the polymer are of approximately the same molecular weight; such polymers are "homodisperse."

[0195] The present invention is further illustrated by reference to a poly(ethylene glycol) or monomethoxy-poly(ethylene glycol) (m-PEG) conjugate. Several reviews and monographs on the functionalization and conjugation of PEG are available. See, for example, Harris, *Macromol. Chem. Phys.* **C25**: 325-373 (1985); Scouten, *Methods in Enzymology* **135**: 30-65 (1987); Wong *et al.*, *Enzyme Microb. Technol.* **14**: 866-874 (1992); Delgado *et al.*, *Critical Reviews in Therapeutic Drug Carrier Systems* **9**: 249-304 (1992); Zalipsky, *Bioconjugate Chem.* **6**: 150-165 (1995); and Bhadra, *et al.*, *Pharmazie*, **57**:5-29 (2002).

[0196] The poly(ethylene glycol) useful in forming the conjugate of the invention is either linear or branched.

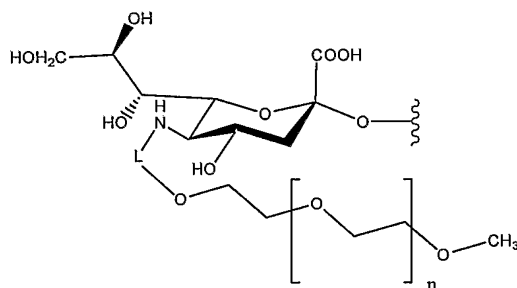
[0197] The *in vivo* half-life of therapeutic glycopeptides can also be enhanced with water-soluble polymers such as polyethylene glycol (PEG, m-PEG) and polypropylene glycol (PPG). For example, chemical modification of proteins with PEG (PEG-ylation, m-PEG-ylation) increases their molecular size and decreases their surface- and functional group-

accessibility, each of which are dependent on the size of the PEG attached to the protein. This results in an improvement of plasma half-lives and in proteolytic-stability, and a decrease in immunogenicity and hepatic uptake (Chaffee *et al. J. Clin. Invest.* **89**: 1643-1651 (1992); Pyatak *et al. Res. Commun. Chem. Pathol Pharmacol.* **29**: 113-127 (1980)).

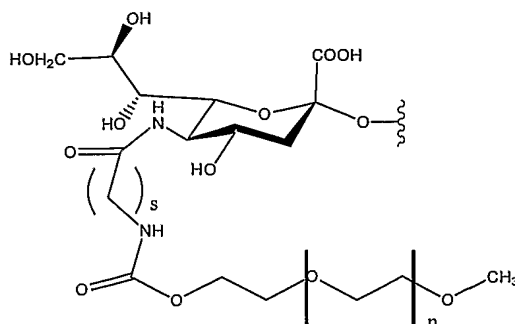
- 5 PEGylation of interleukin-2 has been reported to increase its antitumor potency *in vivo* (Katre *et al. Proc. Natl. Acad. Sci. USA.* **84**: 1487-1491 (1987)) and PEG-ylation of a F(ab')₂ derived from the monoclonal antibody A7 has improved its tumor localization (Kitamura *et al. Biochem. Biophys. Res. Commun.* **28**: 1387-1394 (1990)). Thus, in another preferred embodiment, the *in vivo* half-life of a peptide such as *e.g.*, Glucagon-like peptide-1,
 10 derivatized with a water-soluble polymer by a method of the invention is increased relevant to the *in vivo* half-life of the non-derivatized peptide.

- [0198] The increase in peptide *in vivo* half-life is best expressed as a range of percent increase in this quantity. The lower end of the range of percent increase is about 40%, about 60%, about 80%, about 100%, about 150% or about 200%. The upper end of the range is
 15 about 60%, about 80%, about 100%, about 150%, or more than about 250%.

[0199] In selected glyco-PEG-ylated peptides of the invention, the PEG-intact glycosyl linker cassette has the structure:



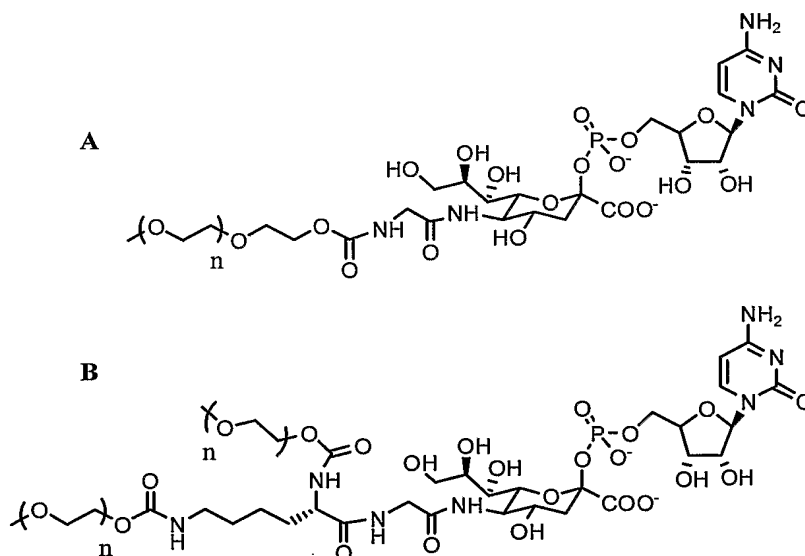
- in which L is a substituted or unsubstituted alkyl or substituted or unsubstituted heteroalkyl
 20 linker moiety joining the sialic acid moiety and the PEG moiety. The index n is selected from the integers from 0 to about 2500, more preferably from about 50 to about 1500, and more preferable still from about 100 to about 600. An example of this structure has the formula:



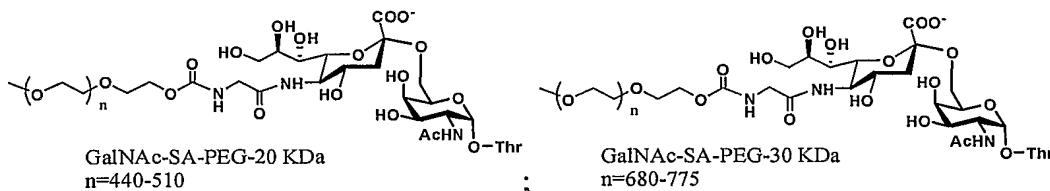
in which the index “s” represents an integer from 0 to 20.

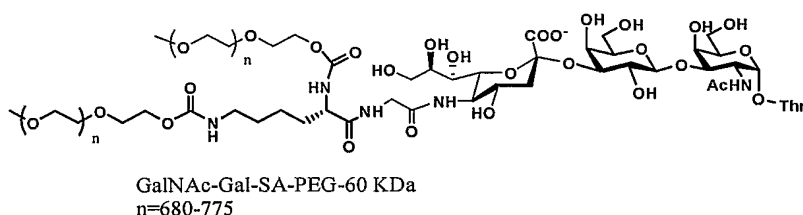
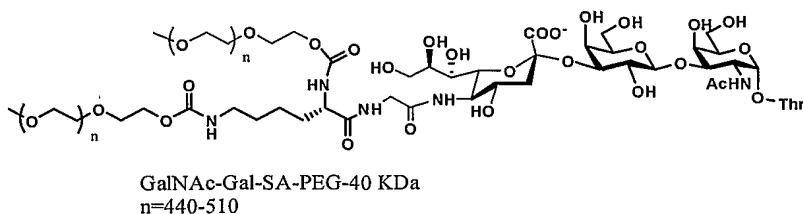
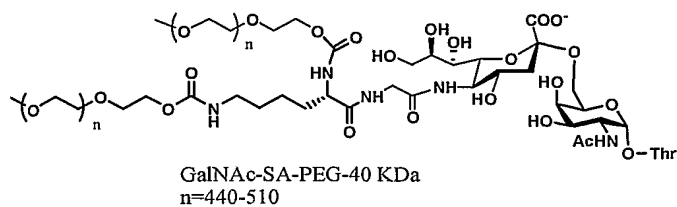
[0200] PEG moieties of any molecular weight, e.g., 5 Kd, 10 Kd, 20 Kd, 30kD, 40 kD, 60 kD and 100kD are of use in the present invention.

- 5 **[0201]** Exemplary activated modified sugars of use in preparing water-soluble polymer-peptide conjugates of the invention include; an linear PEG species (A) and a branched PEG species (B):



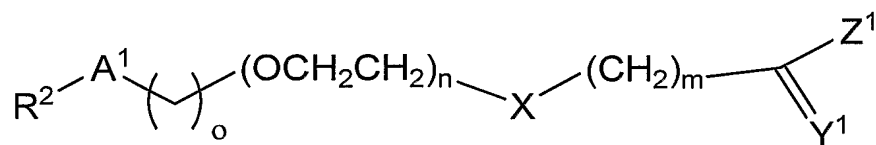
[0202] Following their conjugation to an O-linked site, exemplary PEG-sialic acid-glycosyl moieties can have one or more of the following structures:





In an exemplary embodiment, the Thr shown in the structures above is Thr¹⁰⁶ of interferon alpha 2b.

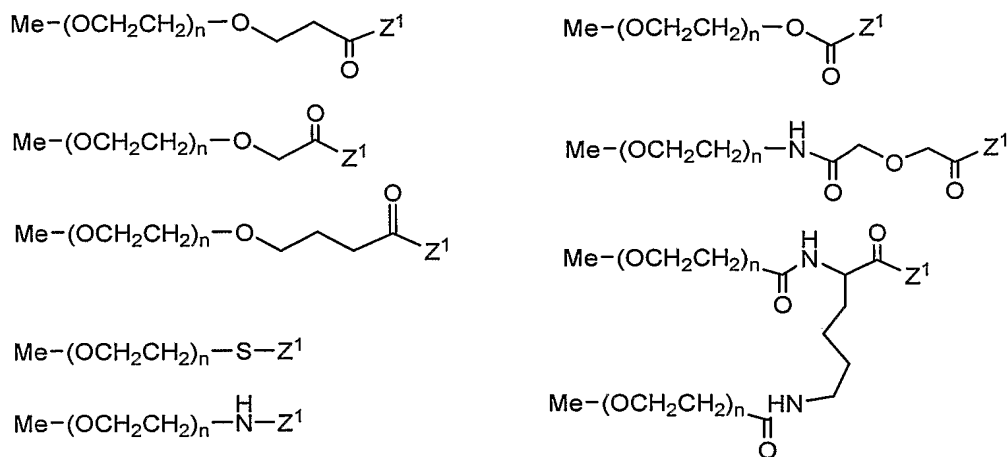
[0203] In another exemplary embodiment, poly(ethylene glycol) molecules of use in the invention include, but are not limited to, those species set forth below.



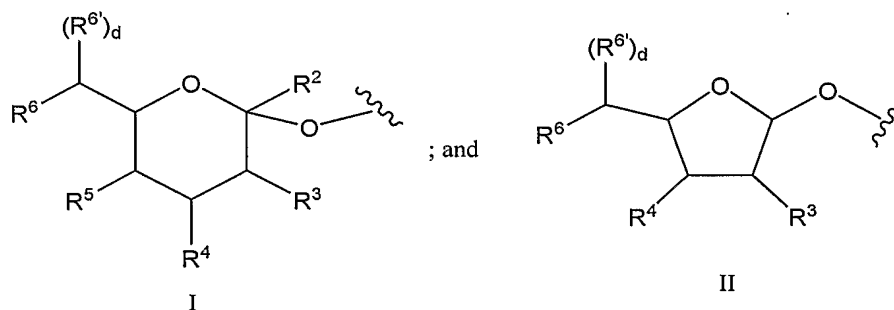
in which R² is H, substituted or unsubstituted alkyl, substituted or unsubstituted aryl, substituted or unsubstituted heteroaryl, substituted or unsubstituted heterocycloalkyl, substituted or unsubstituted heteroalkyl, e.g., acetal, OHC-, H₂N-(CH₂)_q-, HS-(CH₂)_q-, and -(CH₂)_qC(Y¹)Z²; -sugar-nucleotide, or protein. The index “n” represents an integer from 1 to 2500. The indices m, o, and q independently represent integers from 0 to 20. The symbols Z¹ and Z² independently represent OH, NH₂, halogen, S-R³, the alcohol portion of activated esters, -(CH₂)_pC(Y²)V, -(CH₂)_pU(CH₂)_sC(Y²)_v, sugar-nucleotide, protein, and leaving groups, e.g., imidazole, p-nitrophenyl, HOBT, tetrazole, halide. The symbols X, Y¹, Y², A¹, and U independently represent the moieties O, S, N-R⁴. The symbol V represents OH, NH₂, halogen, S-R⁵, the alcohol component of activated esters, the amine component of activated amides, sugar-nucleotides, and proteins. The indices p, q, s and v are members

independently selected from the integers from 0 to 20. The symbols R^3 , R^4 and R^5 independently represent H, substituted or unsubstituted alkyl, substituted or unsubstituted heteroalkyl, substituted or unsubstituted aryl, substituted or unsubstituted heterocycloalkyl and substituted or unsubstituted heteroaryl.

- 5 [0204] In other exemplary embodiments, the poly(ethylene glycol) molecule is selected from the following:



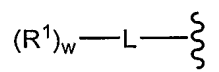
- 10 [0205] In an exemplary embodiment, the invention provides a glycopeptide that is conjugated to a polymeric modifying moiety through an intact glycosyl linking group having a formula that is selected from:



- 15 In Formulae I R^2 is H, CH_2OR^7 , COOR^7 or OR^7 , in which R^7 represents H, substituted or unsubstituted alkyl or substituted or unsubstituted heteroalkyl. When COOR^7 is a carboxylic acid or carboxylate, both forms are represented by the designation of the single structure COO^- or COOH . In Formulae I and II, the symbols R^3 , R^4 , R^5 , R^6 and $R^{6'}$ independently represent H, substituted or unsubstituted alkyl, OR^8 , NHC(O)R^9 . The index d is 0 or 1. R^8 and R^9 are independently selected from H, substituted or unsubstituted alkyl, substituted or unsubstituted heteroalkyl, sialic acid or polysialic acid. At least one of R^3 , R^4 , R^5 , R^6 or $R^{6'}$

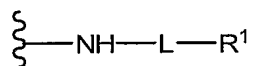
includes the polymeric modifying moiety *e.g.*, PEG, linked through a bond or a linking group. In an exemplary embodiment, R⁶ and R^{6'}, together with the carbon to which they are attached are components of the pyruvyl side chain of sialic acid. In a further exemplary embodiment, this side chain is functionalized with the polymeric modifying moiety. In another exemplary embodiment, R⁶ and R^{6'}, together with the carbon to which they are attached are components of the side chain of sialic acid and the polymeric modifying moiety is a component of R⁵.

[0206] In a further exemplary embodiment, the polymeric modifying moiety is bound to the sugar core, generally through a heteroatom, *e.g.*, nitrogen, on the core through a linker, L, as shown below:

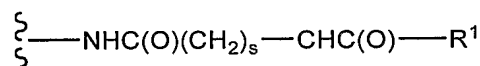


R¹ is the polymeric moiety and L is selected from a bond and a linking group. The index w represents an integer selected from 1-6, preferably 1-3 and more preferably 1-2. Exemplary linking groups include substituted or unsubstituted alkyl, substituted or unsubstituted heteroalkyl moieties and sialic acid. An exemplary component of the linker is an acyl moiety.

[0207] An exemplary compound according to the invention has a structure according to Formulae I or II, in which at least one of R², R³, R⁴, R⁵, R⁶ or R^{6'} has the formula:

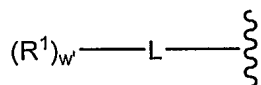


[0208] In another example according to this embodiment at least one of R², R³, R⁴, R⁵, R⁶ or R^{6'} has the formula:



in which s is an integer from 0 to 20 and R¹ is a linear polymeric modifying moiety.

[0209] In an exemplary embodiment, the polymeric modifying moiety-linker construct is a branched structure that includes two or more polymeric chains attached to the central moiety. In this embodiment, the construct has the formula:



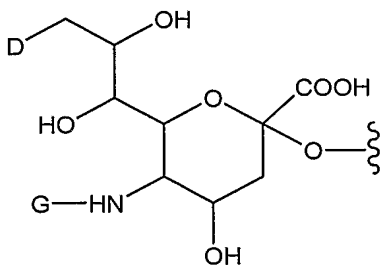
in which R^1 and L are as discussed above and w' is an integer from 2 to 6, preferably from 2 to 4 and more preferably from 2 to 3.

[0210] When L is a bond it is formed between a reactive functional group on a precursor of R^1 and a reactive functional group of complementary reactivity on the saccharyl core. When L is a non-zero order linker, a precursor of L can be in place on the glycosyl moiety prior to reaction with the R^1 precursor. Alternatively, the precursors of R^1 and L can be incorporated into a preformed cassette that is subsequently attached to the glycosyl moiety. As set forth herein, the selection and preparation of precursors with appropriate reactive functional groups is within the ability of those skilled in the art. Moreover, coupling the precursors proceeds by chemistry that is well understood in the art.

[0211] In an exemplary embodiment, L is a linking group that is formed from an amino acid, or small peptide (e.g., 1-4 amino acid residues) providing a modified sugar in which the polymeric modifying moiety is attached through a substituted alkyl linker. Exemplary linkers include glycine, lysine, serine and cysteine. The PEG moiety can be attached to the amine moiety of the linker through an amide or urethane bond. The PEG is linked to the sulfur or oxygen atoms of cysteine and serine through thioether or ether bonds, respectively.

[0212] In an exemplary embodiment, R^5 includes the polymeric modifying moiety. In another exemplary embodiment, R^5 includes both the polymeric modifying moiety and a linker, L, joining the modifying moiety to the remainder of the molecule. As discussed above, L can be a linear or branched structure. Similarly, the polymeric modifying can be branched or linear.

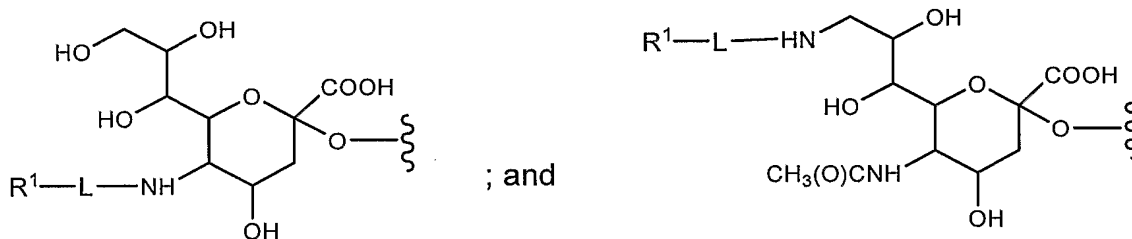
[0213] In one embodiment, the present invention provides an GLP-1 peptide comprising the moiety:



wherein D is a member selected from -OH and R^1 -L-HN-; G is a member selected from H and R^1 -L- and -C(O)(C_1 - C_6)alkyl; R^1 is a moiety comprising a straight-chain or branched poly(ethylene glycol) residue; and L is a linker, e.g., a bond ("zero order"), substituted or

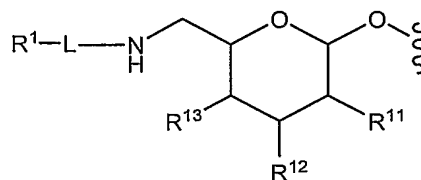
unsubstituted alkyl and substituted or unsubstituted heteroalkyl. In exemplary embodiments, when D is OH, G is R^1 -L-, and when G is $-C(O)(C_1-C_6)\text{alkyl}$, D is R^1 -L-NH-.

[0214] In another exemplary embodiment, the invention provides a conjugate formed between a modified sugar of the invention and a substrate GLP-1 peptide. In this embodiment, the sugar moiety of the modified sugar becomes a glycosyl linking group interposed between the peptide substrate and the modifying group. An exemplary glycosyl linking group is an intact glycosyl linking group, in which the glycosyl moiety or moieties forming the linking group are not degraded by chemical (e.g., sodium metaperiodate) or enzymatic (e.g., oxidase) processes. Selected conjugates of the invention include a modifying group that is attached to the amine moiety of an amino-saccharide, e.g., mannosamine, glucosamine, galactosamine, sialic acid etc. Exemplary modifying group-intact glycosyl linking group cassettes according to this motif are based on a sialic acid structure, such as those having the formulae:



[0215] In the formulae above, R^1 and L are as described above. Further detail about the structure of exemplary R^1 groups is provided below.

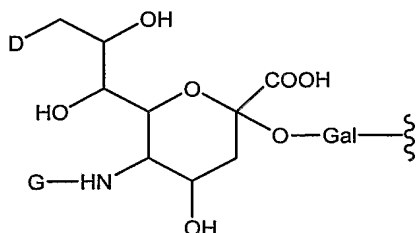
[0216] In still a further exemplary embodiment, the conjugate is formed between a substrate GLP-1 and a saccharyl moiety in which the modifying group is attached through a linker at the 6-carbon position of the saccharyl moiety. Thus, illustrative conjugates according to this embodiment have the formula:



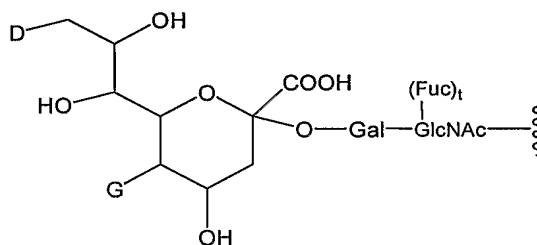
in which the radicals are as discussed above. Such saccharyl moieties include, without limitation, glucose, glucosamine, N-acetyl-glucosamine, galactose, galactosamine, N-acetyl-galactosamine, mannose, mannosamine, N-acetyl-mannosamine, and the like.

[0217] Due to the versatility of the methods available for modifying glycosyl residues on a therapeutic peptide such as GLP-1, the glycosyl structures on the peptide conjugates of the invention can have substantially any structure. Moreover, the glycans can be O-linked or N-linked. As exemplified in the discussion below, each of the pyranose and furanose derivatives discussed above can be a component of a glycosyl moiety of a peptide.

[0218] The invention provides a modified GLP-1 peptide that includes a glycosyl group having the formula:

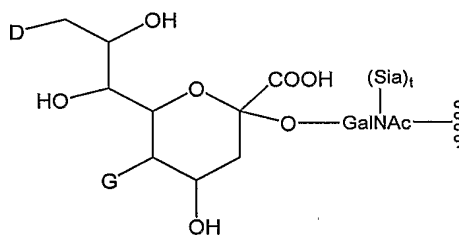


[0219] In other embodiments, the group has the formula:



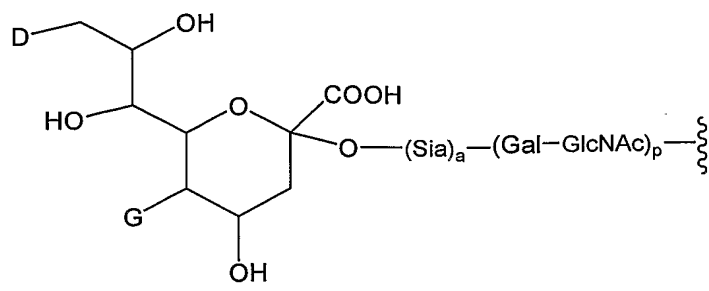
in which the index t is 0 or 1.

[0220] In a still further exemplary embodiment, the group has the formula:



in which the index t is 0 or 1.

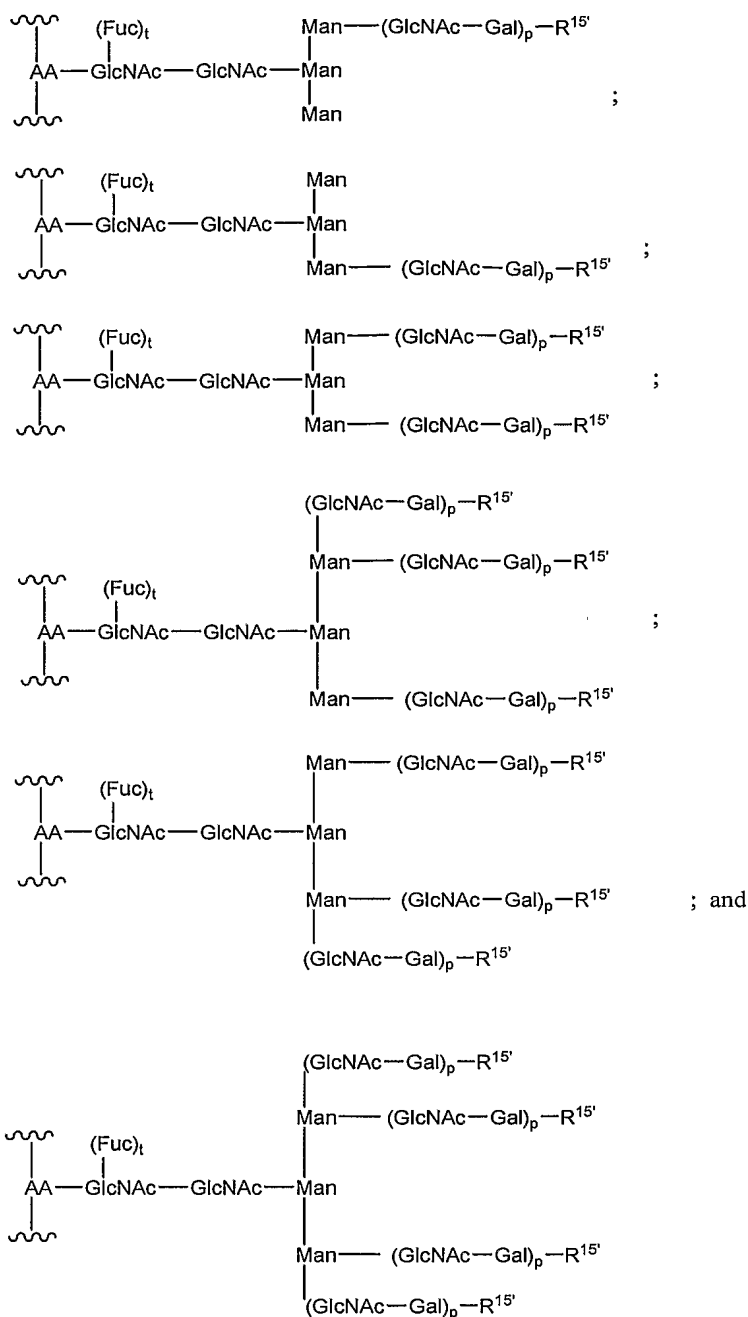
[0221] In yet another embodiment, the group has the formula:



in which the index p represents an integer from 1 to 10; and a is either 0 or 1.

[0222] In an exemplary embodiment, a glycoPEGylated GLP-1 peptide of the invention includes at least one N-linked glycosyl residue selected from the glycosyl residues set forth

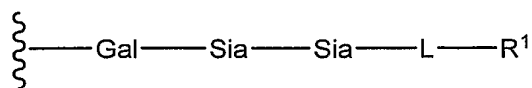
5 below:



[0223] In the formulae above, the index t is 0 or 1 and the index p is an integer from 1 to 10. The symbol $\text{R}^{15'}$ represents H, OH (e.g., Gal-OH), a sialyl moiety, a polymer modified sialyl moiety (i.e., glycosyl linking group-polymeric modifying moiety (Sia-L-R^1)) or a sialyl moiety to which is bound a polymer modified sialyl moiety (e.g., Sia-Sia-L-R^1) (“Sia-Sia^p”). Exemplary polymer modified saccharyl moieties have a structure according to Formulae I and II. An exemplary GLP-1 peptide of the invention will include at least one glycan having a $\text{R}^{15'}$ that includes a structure according to Formulae I or II. The oxygen, with the open

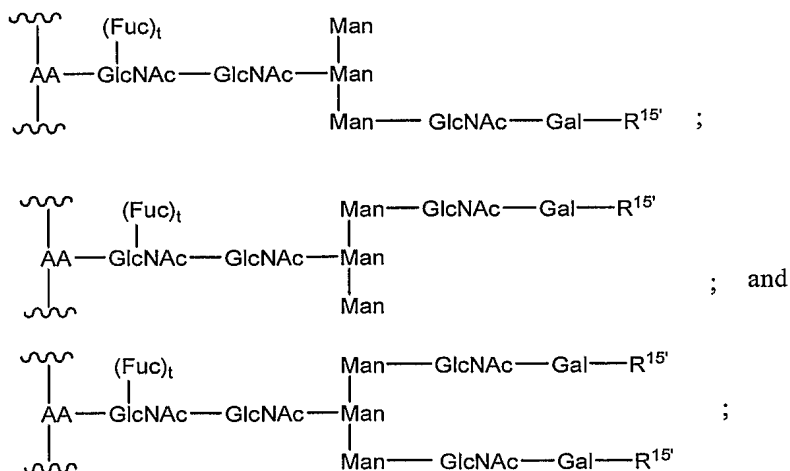
valence, of Formulae I and II is preferably attached through a glycosidic linkage to a carbon of a Gal or GalNAc moiety. In a further exemplary embodiment, the oxygen is attached to the carbon at position 3 of a galactose residue. In an exemplary embodiment, the modified sialic acid is linked α 2,3-to the galactose residue. In another exemplary embodiment, the sialic acid is linked α 2,6-to the galactose residue.

[0224] In another exemplary embodiment, the invention provides a GLP-1 peptide conjugate that includes a glycosyl linking group, such as those set forth above, that is covalently attached to an amino acid residue of the peptide. In one embodiment according to this motif, the glycosyl linking moiety is linked to a galactose residue through a Sia residue:



An exemplary species according to this motif is prepared by conjugating Sia-L-R¹ to a terminal sialic acid of a glycan using an enzyme that forms Sia-Sia bonds, e.g., CST-II, ST8Sia-II, ST8Sia-III and ST8Sia-IV.

[0225] In another exemplary embodiment, the glycans have a formula that is selected from the group:



and combinations thereof.

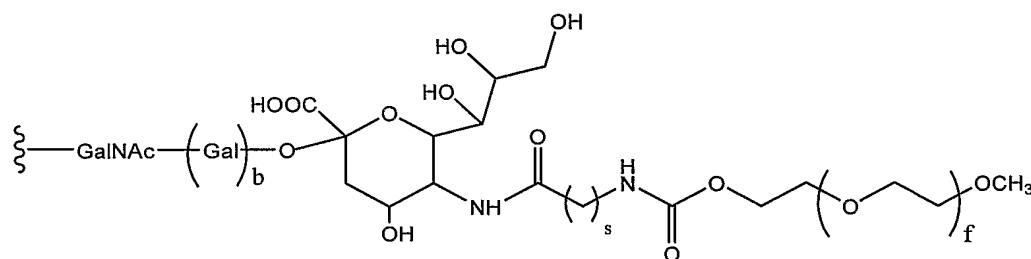
[0226] The glycans of this group generally correspond to those found on a GLP-1 peptide that is produced by insect (e.g., Sf-9) cells, following remodeling according to the methods set forth herein. For example insect-derived GLP-1 that is expressed with a tri-mannosyl core is subsequently contacted with a GlcNAc donor and a GlcNAc transferase and a Gal

donor and a Gal transferase. Appending GlcNAc and Gal to the tri-mannosyl core is accomplished in either two steps or a single step. A modified sialic acid is added to at least one branch of the glycosyl moiety as discussed herein. Those Gal moieties that are not functionalized with the modified sialic acid are optionally "capped" by reaction with a sialic acid donor in the presence of a sialyl transferase.

[0227] In an exemplary embodiment, at least 60% of terminal Gal moieties in a population of peptides is capped with sialic acid, preferably at least 70%, more preferably, at least 80%, still more preferably at least 90% and even more preferably at least 95%, 96%, 97%, 98% or 99% are capped with sialic acid.

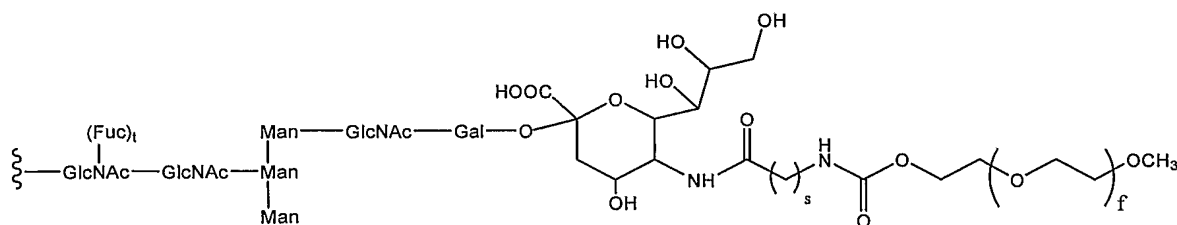
10 [0228] In each of the formulae above, $R^{15'}$ is as discussed above. Moreover, an exemplary modified GLP-1 peptide of the invention will include at least one glycan with an R^{15} moiety having a structure according to Formulae I or II.

[0229] In an exemplary embodiment, the glycosyl linking moiety has the formula:



15 in which b is 0 or 1. The index s represents an integer from 1 to 10; and f represents an integer from 1 to 2500. Generally preferred is the use of a PEG moiety that has a molecular weight of about 20 kDa.

[0230] In another exemplary embodiment, the GLP-1 is derived from insect cells, remodeled by adding GlcNAc and Gal to the mannose core and glycopegylated using a sialic acid bearing a linear PEG moiety, affording a GLP-1 peptide that comprises at least one moiety having the formula:



in which s represents an integer from 1 to 10; and f represents an integer from 1 to 2500.

[0231] As discussed herein, the PEG of use in the conjugates of the invention can be linear or branched. An exemplary precursor of use to form the branched conjugates according to this embodiment of the invention has the formula:



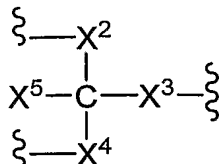
[0232] The branched polymer species according to this formula are essentially pure water-soluble polymers. $X^{3'}$ is a moiety that includes an ionizable, e.g., OH, COOH, H_2PO_4 , HSO_3 , HPO_3 , and salts thereof, etc.) or other reactive functional group, e.g., *infra*. C is carbon. X^5 is preferably a non-reactive group (e.g., H, unsubstituted alkyl, unsubstituted heteroalkyl), and can be a polymeric arm. R^{16} and R^{17} are independently selected polymeric arms, e.g., nonpeptidic, nonreactive polymeric arms (e.g., PEG)). X^2 and X^4 are linkage fragments that are preferably essentially non-reactive under physiological conditions, which may be the same or different. An exemplary linker includes neither aromatic nor ester moieties. Alternatively, these linkages can include one or more moiety that is designed to degrade under physiologically relevant conditions, e.g., esters, disulfides, etc. X^2 and X^4 join polymeric arms R^{16} and R^{17} to C. When $X^{3'}$ is reacted with a reactive functional group of complementary reactivity on a linker, sugar or linker-sugar cassette, $X^{3'}$ is converted to a component of linkage fragment X^3 .

[0233] Exemplary linkage fragments for X^2 , X^3 and X^4 are independently selected and include S, SC(O)NH, HNC(O)S, SC(O)O, O, NH, NHC(O), (O)CNH and NHC(O)O, and OC(O)NH, CH_2S , CH_2O , CH_2CH_2O , CH_2CH_2S , $(CH_2)_oO$, $(CH_2)_oS$ or $(CH_2)_oY'-PEG$ wherein, Y' is S, NH, NHC(O), C(O)NH, NHC(O)O, OC(O)NH, or O and o is an integer from 1 to 50. In an exemplary embodiment, the linkage fragments X^2 and X^4 are different linkage fragments.

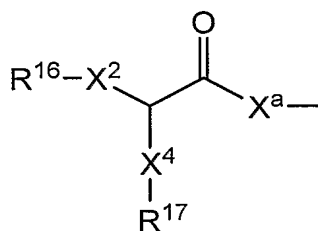
[0234] In an exemplary embodiment, the precursor (III), or an activated derivative thereof, is reacted with, and thereby bound to a sugar, an activated sugar or a sugar nucleotide through a reaction between $X^{3'}$ and a group of complementary reactivity on the sugar moiety, e.g., an amine. Alternatively, $X^{3'}$ reacts with a reactive functional group on a precursor to linker, L.

One or more of R^2 , R^3 , R^4 , R^5 , R^6 or R^6' of Formulae I and II can include the branched polymeric modifying moiety, or this moiety bound through L.

[0235] In an exemplary embodiment, the moiety:



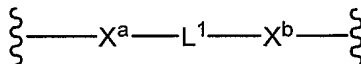
- 5 is the linker arm, L. In this embodiment, an exemplary linker is derived from a natural or unnatural amino acid, amino acid analogue or amino acid mimetic, or a small peptide formed from one or more such species. For example, certain branched polymers found in the compounds of the invention have the formula:



(IV)

- 10 [0236] X^a is a linkage fragment that is formed by the reaction of a reactive functional group, e.g., $X^{3'}$, on a precursor of the branched polymeric modifying moiety and a reactive functional group on the sugar moiety, or a precursor to a linker. For example, when $X^{3'}$ is a carboxylic acid, it can be activated and bound directly to an amine group pendent from an amino-saccharide (e.g., Sia, GalNH₂, GlcNH₂, ManNH₂, etc.), forming an X^a that is an
- 15 amide. Additional exemplary reactive functional groups and activated precursors are described hereinbelow. The index c represents an integer from 1 to 10. The other symbols have the same identity as those discussed above.

[0237] In another exemplary embodiment, X^a is a linking moiety formed with another linker:



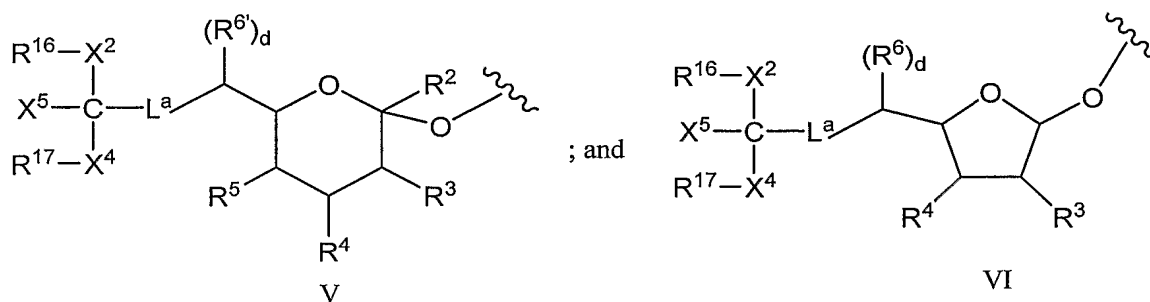
20

in which X^b is a second linkage fragment and is independently selected from those groups set forth for X^a , and, similar to L, L^1 is a bond, substituted or unsubstituted alkyl or substituted or unsubstituted heteroalkyl.

[0238] Exemplary species for X^a and X^b include S, SC(O)NH, HNC(O)S, SC(O)O, O, NH, NHC(O), C(O)NH and NHC(O)O, and OC(O)NH.

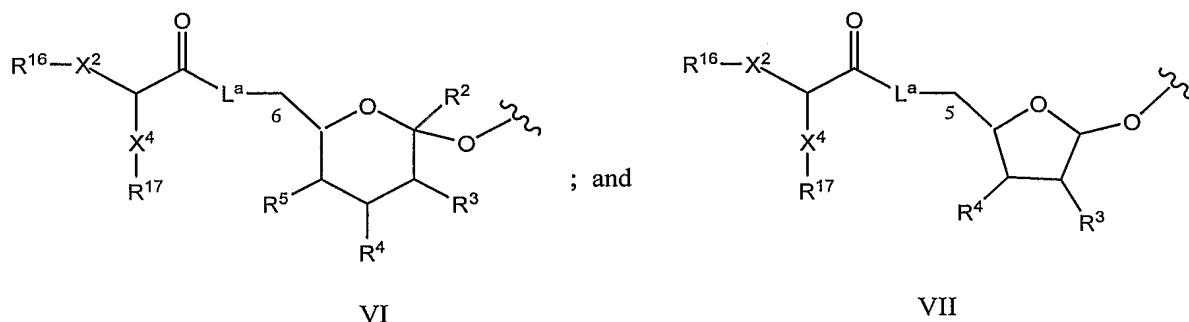
[0239] In another exemplary embodiment, X^4 is a peptide bond to R^{17} , which is an amino acid, di-peptide (e.g., Lys-Lys) or tri-peptide (E.G., Lys-Lys-Lys) in which the alpha-amine moiety(ies) and/or side chain heteroatom(s) are modified with a polymeric modifying moiety.

[0240] In a further exemplary embodiment, the conjugates of the invention include a moiety, e.g., an R^{15} moiety that has a formula that is selected from:



in which the identity of the radicals represented by the various symbols is the same as that discussed hereinabove. L^a is a bond or a linker as discussed above for L and L^1 , e.g., substituted or unsubstituted alkyl or substituted or unsubstituted heteroalkyl moiety. In an exemplary embodiment, L^a is a moiety of the side chain of sialic acid that is functionalized with the polymeric modifying moiety as shown. Exemplary L^a moieties include substituted or unsubstituted alkyl chains that include one or more OH or NH_2 .

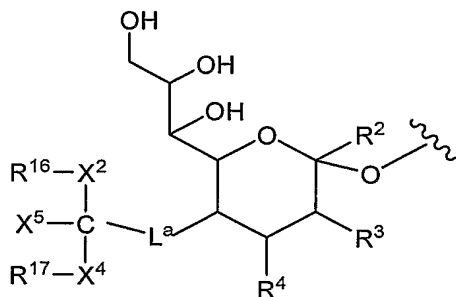
[0241] In yet another exemplary embodiment, the invention provides conjugates having a moiety, e.g., an R^{15} moiety with formula:



The identity of the radicals represented by the various symbols is the same as that discussed hereinabove. As those of skill will appreciate, the linker arm in Formulae VI and VII is equally applicable to other modified sugars set forth herein. In exemplary embodiment, the

species of Formulae VI and VII are the R^{15} moieties attached to the glycan structures set forth herein.

[0242] In yet another exemplary embodiment, the GLP-1 peptide includes an R^{15} moiety with the formula:

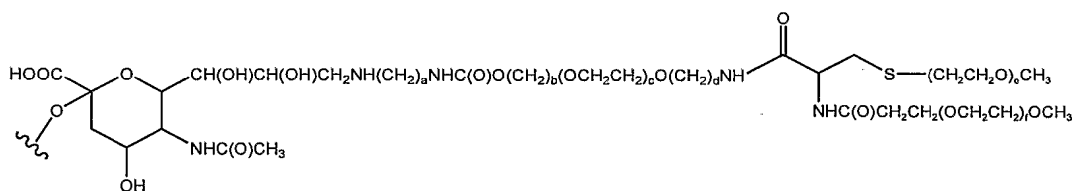
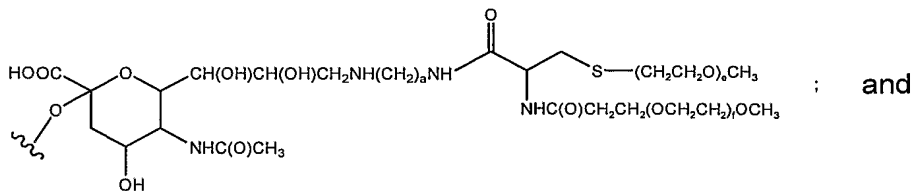
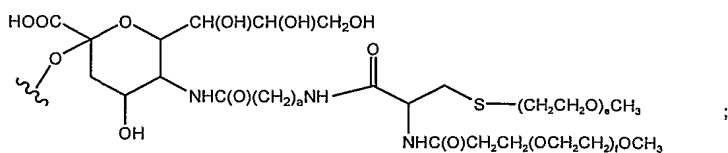
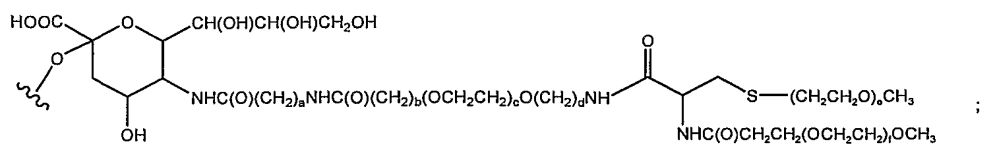


in which the identities of the radicals are as discussed above. An exemplary species for L^a is $-(CH_2)_jC(O)NH(CH_2)_hC(O)NH-$, in which h and j are independently selected integers from 0 to 10. A further exemplary species is $-C(O)NH-$.

[0243] The embodiments of the invention set forth above are further exemplified by reference to species in which the polymer is a water-soluble polymer, particularly poly(ethylene glycol) ("PEG"), e.g., methoxy-poly(ethylene glycol). Those of skill will appreciate that the focus in the sections that follow is for clarity of illustration and the various motifs set forth using PEG as an exemplary polymer are equally applicable to species in which a polymer other than PEG is utilized.

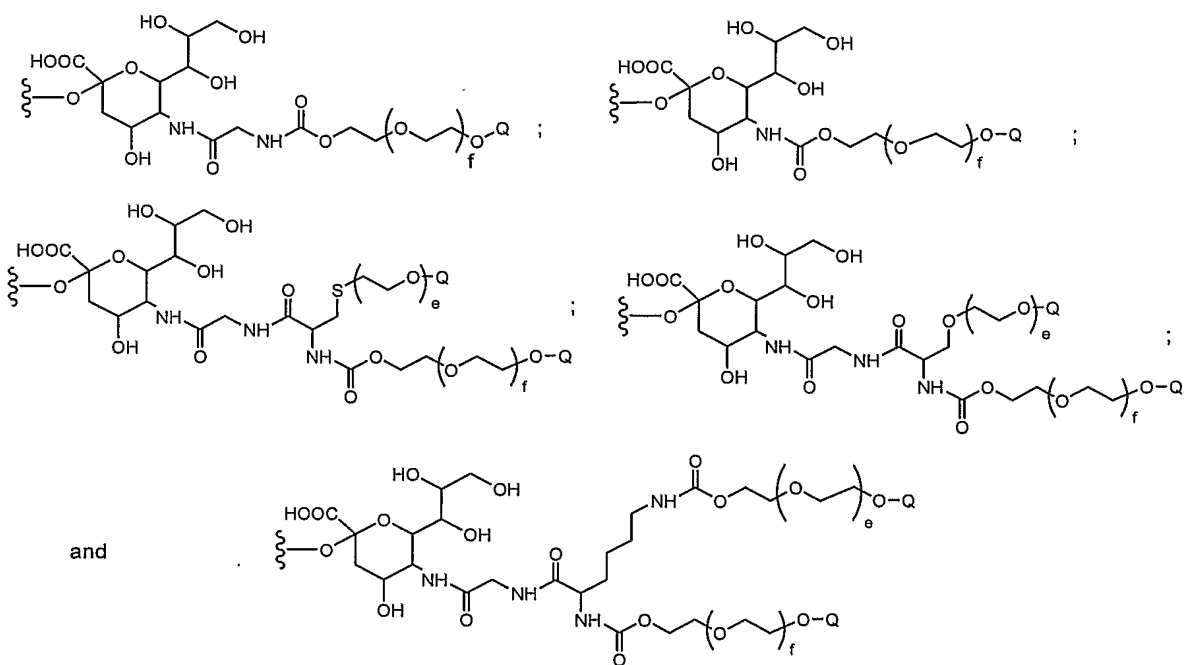
[0244] PEG of any molecular weight, e.g., 1 kDa, 2 kDa, 5 kDa, 10 kDa, 15 kDa, 20 kDa, 30 kDa and 40 kDa is of use in the present invention.

[0245] In an exemplary embodiment, the R^{15} moiety has a formula that is a member selected from the group:



In each of the structures above, the linker fragment $-\text{NH}(\text{CH}_2)_a-$ can be present or absent.

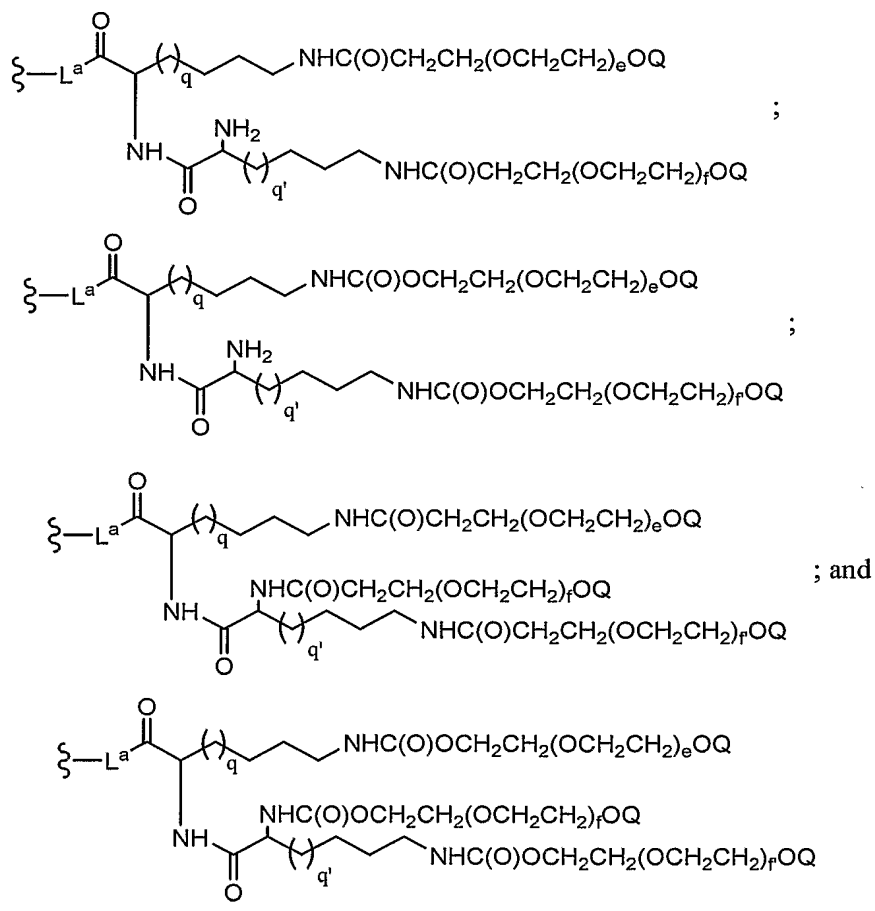
[0246] In other exemplary embodiments, the conjugate includes an R^{15} moiety selected from the group:



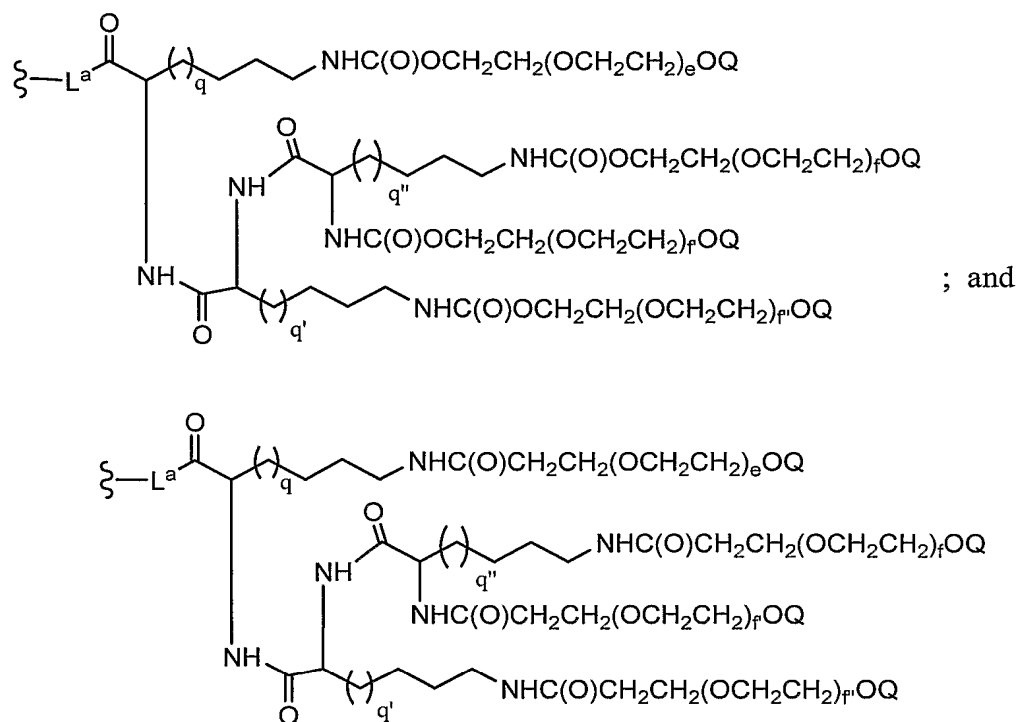
[0247] In each of the formulae above, the indices e and f are independently selected from the integers from 1 to 2500. In further exemplary embodiments, e and f are selected to provide a PEG moiety that is about 1 kD, 2 kD, 10 kD, 15 kD, 20 kD, 30 kD or 40 kD. The symbol Q represents substituted or unsubstituted alkyl (e.g., C₁-C₆ alkyl, e.g., methyl),

5 substituted or unsubstituted heteroalkyl or H.

[0248] Other branched polymers have structures based on di-lysine (Lys-Lys) peptides, e.g.:

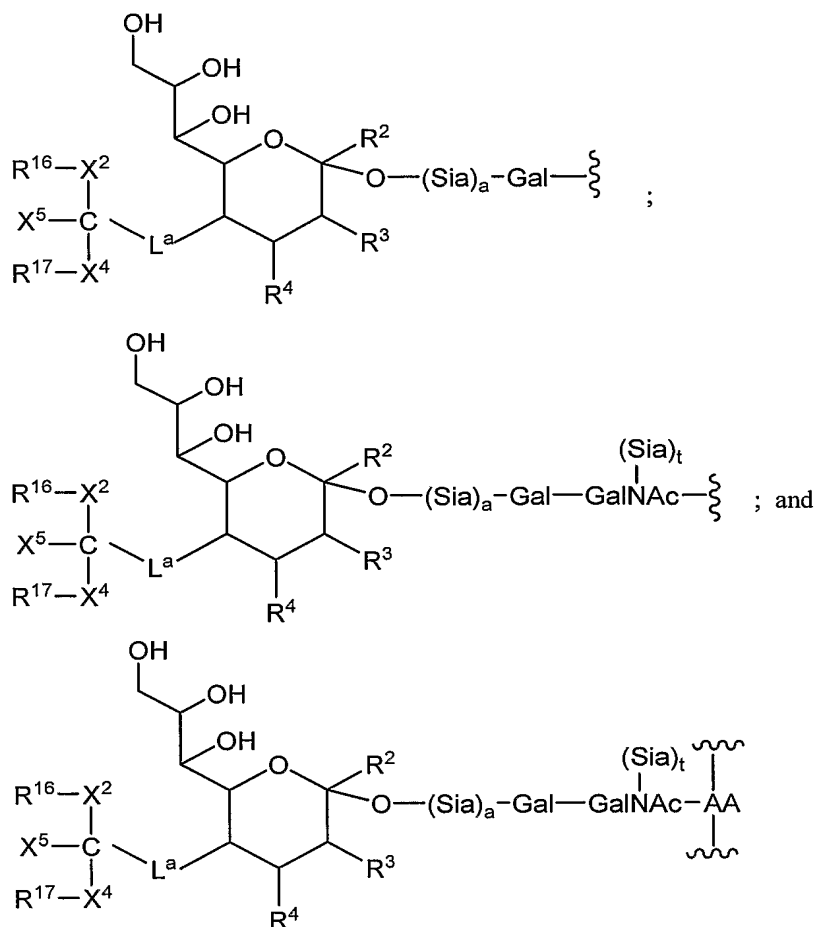


and tri-lysine peptides (Lys-Lys-Lys), e.g.:



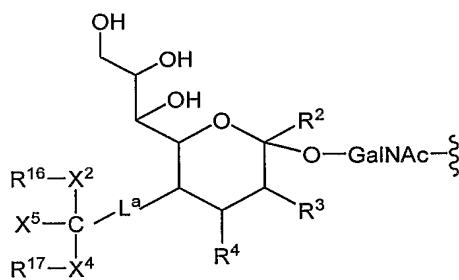
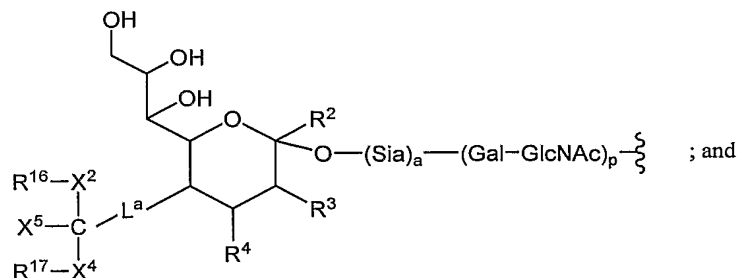
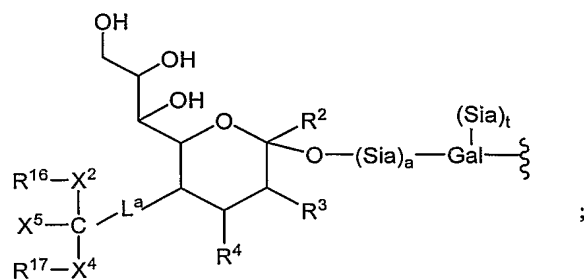
In each of the figures above, e, f, f' and f'' represent integers independently selected from 1 to 2500. The indices q, q' and q'' represent integers independently selected from 1 to 20.

[0249] In another exemplary embodiment, the GLP-1 peptide comprises a glycosyl moiety selected from the formulae:



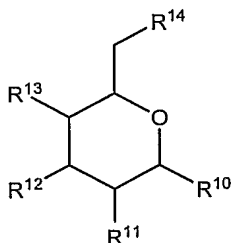
in which L^a is a bond or a linker as described herein; the index t represents 0 or 1; and the index a represents 0 or 1. Each of these groups can be included as components of the mono-, bi-, tri- and tetra-antennary saccharide structures set forth above.

- 5 **[0250]** In yet another embodiment, the conjugates of the invention include a modified glycosyl residue that includes the substructure selected from:



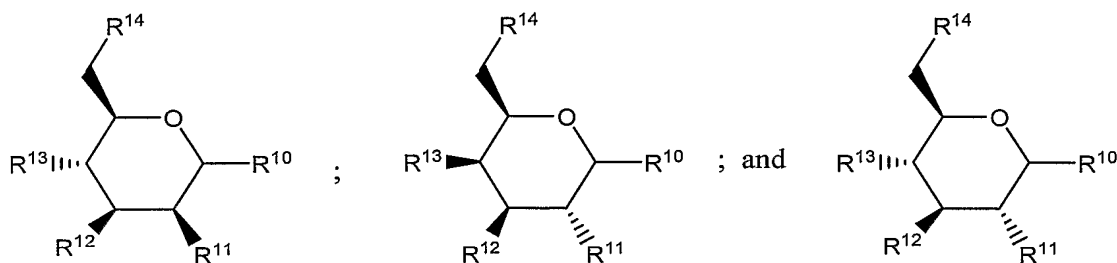
in which the index a and the linker L^a are as discussed above. The index p is an integer from 1 to 10. The indices t and a are independently selected from 0 or 1. Each of these groups can be included as components of the mono-, bi-, tri- and tetra-antennary saccharide structures set forth above.

[0251] In a further exemplary embodiment, the invention utilizes modified sugars in which the 6-hydroxyl position is converted to the corresponding amine moiety, which bears a linker-modifying group cassette such as those set forth above. Exemplary saccharyl groups that can be used as the core of these modified sugars include Gal, GalNAc, Glc, GlcNAc, Fuc, Xyl, Man, and the like. A representative modified sugar according to this embodiment has the formula:



in which R^{11} - R^{14} are members independently selected from H, OH, $C(O)CH_3$, NH, and $NH-C(O)CH_3$. R^{10} is a link to another glycosyl residue (-O-glycosyl) or to an amino acid of the GLP-1 peptide (-NH-(GLP-1)). R^{14} is OR^1 , NHR^1 or $NH-L-R^1$. R^1 and $NH-L-R^1$ are as described above.

- 5 [0252] Selected conjugates according to this motif are based on mannose, galactose or glucose, or on species having the stereochemistry of mannose, galactose or glucose. The general formulae of these conjugates are:



- 10 [0253] As discussed above, the invention provides saccharides bearing a modifying group, activated analogues of these species and conjugates formed between species such as peptides and lipids and a modified saccharide of the invention.

Biomolecules

- 15 [0254] In another preferred embodiment, the modified sugar bears a biomolecule. In still further preferred embodiments, the biomolecule is a functional protein, enzyme, antigen, antibody, peptide, nucleic acid (*e.g.*, single nucleotides or nucleosides, oligonucleotides, polynucleotides and single- and higher-stranded nucleic acids), lectin, receptor or a combination thereof.

- 20 [0255] Preferred biomolecules are essentially non-fluorescent, or emit such a minimal amount of fluorescence that they are inappropriate for use as a fluorescent marker in an assay. Moreover, it is generally preferred to use biomolecules that are not sugars. An exception to this preference is the use of an otherwise naturally occurring sugar that is modified by covalent attachment of another entity (*e.g.*, PEG, biomolecule, therapeutic moiety, diagnostic moiety, *etc.*). In an exemplary embodiment, a sugar moiety, which is a biomolecule, is conjugated to a linker arm and the sugar-linker arm cassette is subsequently conjugated to a peptide via a method of the invention.
- 25

[0256] Biomolecules useful in practicing the present invention can be derived from any source. The biomolecules can be isolated from natural sources or they can be produced by synthetic methods. Peptides can be natural peptides or mutated peptides. Mutations can be effected by chemical mutagenesis, site-directed mutagenesis or other means of inducing mutations known to those of skill in the art. Peptides useful in practicing the instant invention include, for example, enzymes, antigens, antibodies and receptors. Antibodies can be either polyclonal or monoclonal; either intact or fragments. The peptides are optionally the products of a program of directed evolution

[0257] Both naturally derived and synthetic peptides and nucleic acids are of use in conjunction with the present invention; these molecules can be attached to a sugar residue component or a crosslinking agent by any available reactive group. For example, peptides can be attached through a reactive amine, carboxyl, sulfhydryl, or hydroxyl group. The reactive group can reside at a peptide terminus or at a site internal to the peptide chain. Nucleic acids can be attached through a reactive group on a base (*e.g.*, exocyclic amine) or an available hydroxyl group on a sugar moiety (*e.g.*, 3'- or 5'-hydroxyl). The peptide and nucleic acid chains can be further derivatized at one or more sites to allow for the attachment of appropriate reactive groups onto the chain. *See, Chrisey et al. Nucleic Acids Res.* **24**: 3031-3039 (1996).

[0258] In a further preferred embodiment, the biomolecule is selected to direct the peptide modified by the methods of the invention to a specific tissue, thereby enhancing the delivery of the peptide to that tissue relative to the amount of underivatized peptide that is delivered to the tissue. In a still further preferred embodiment, the amount of derivatized peptide delivered to a specific tissue within a selected time period is enhanced by derivatization by at least about 20%, more preferably, at least about 40%, and more preferably still, at least about 100%. Presently, preferred biomolecules for targeting applications include antibodies, hormones and ligands for cell-surface receptors.

[0259] In still a further exemplary embodiment, there is provided as conjugate with biotin. Thus, for example, a selectively biotinylated peptide is elaborated by the attachment of an avidin or streptavidin moiety bearing one or more modifying groups.

Therapeutic Moieties

[0260] In another preferred embodiment, the modified sugar includes a therapeutic moiety. Those of skill in the art will appreciate that there is overlap between the category of

therapeutic moieties and biomolecules; many biomolecules have therapeutic properties or potential.

[0261] The therapeutic moieties can be agents already accepted for clinical use or they can be drugs whose use is experimental, or whose activity or mechanism of action is under investigation. The therapeutic moieties can have a proven action in a given disease state or can be only hypothesized to show desirable action in a given disease state. In a preferred embodiment, the therapeutic moieties are compounds, which are being screened for their ability to interact with a tissue of choice. Therapeutic moieties, which are useful in practicing the instant invention include drugs from a broad range of drug classes having a variety of pharmacological activities. Preferred therapeutic moieties are essentially non-fluorescent, or emit such a minimal amount of fluorescence that they are inappropriate for use as a fluorescent marker in an assay. Moreover, it is generally preferred to use therapeutic moieties that are not sugars. An exception to this preference is the use of a sugar that is modified by covalent attachment of another entity, such as a PEG, biomolecule, therapeutic moiety, diagnostic moiety and the like. In another exemplary embodiment, a therapeutic sugar moiety is conjugated to a linker arm and the sugar-linker arm cassette is subsequently conjugated to a peptide via a method of the invention.

[0262] Methods of conjugating therapeutic and diagnostic agents to various other species are well known to those of skill in the art. See, for example Hermanson, BIOCONJUGATE TECHNIQUES, Academic Press, San Diego, 1996; and Dunn *et al.*, Eds. POLYMERIC DRUGS AND DRUG DELIVERY SYSTEMS, ACS Symposium Series Vol. 469, American Chemical Society, Washington, D.C. 1991.

[0263] In an exemplary embodiment, the therapeutic moiety is attached to the modified sugar via a linkage that is cleaved under selected conditions. Exemplary conditions include, but are not limited to, a selected pH (e.g., stomach, intestine, endocytotic vacuole), the presence of an active enzyme (e.g, esterase, reductase, oxidase), light, heat and the like. Many cleavable groups are known in the art. See, for example, Jung *et al.*, *Biochem. Biophys. Acta*, **761**: 152-162 (1983); Joshi *et al.*, *J. Biol. Chem.*, **265**: 14518-14525 (1990); Zarling *et al.*, *J. Immunol.*, **124**: 913-920 (1980); Bouizar *et al.*, *Eur. J. Biochem.*, **155**: 141-147 (1986); Park *et al.*, *J. Biol. Chem.*, **261**: 205-210 (1986); Browning *et al.*, *J. Immunol.*, **143**: 1859-1867 (1989).

Preparation of Modified Sugars

[0264] In general, the sugar moiety and the modifying group are linked together through the use of reactive groups, which are typically transformed by the linking process into a new organic functional group or unreactive species. The sugar reactive functional group(s), is located at any position on the sugar moiety. Reactive groups and classes of reactions useful in practicing the present invention are generally those that are well known in the art of bioconjugate chemistry. Currently favored classes of reactions available with reactive sugar moieties are those, which proceed under relatively mild conditions. These include, but are not limited to nucleophilic substitutions (*e.g.*, reactions of amines and alcohols with acyl halides, active esters), electrophilic substitutions (*e.g.*, enamine reactions) and additions to carbon-carbon and carbon-heteroatom multiple bonds (*e.g.*, Michael reaction, Diels-Alder addition). These and other useful reactions are discussed in, for example, March, *ADVANCED ORGANIC CHEMISTRY*, 3rd Ed., John Wiley & Sons, New York, 1985; Hermanson, *BIOCONJUGATE TECHNIQUES*, Academic Press, San Diego, 1996; and Feeney *et al.*, *MODIFICATION OF PROTEINS*; *Advances in Chemistry Series*, Vol. 198, American Chemical Society, Washington, D.C., 1982.

[0265] Useful reactive functional groups pendent from a sugar nucleus or modifying group include, but are not limited to:

- (a) carboxyl groups and various derivatives thereof including, but not limited to, N-hydroxysuccinimide esters, N-hydroxybenztriazole esters, acid halides, acyl imidazoles, thioesters, p-nitrophenyl esters, alkyl, alkenyl, alkynyl and aromatic esters;
- (b) hydroxyl groups, which can be converted to, *e.g.*, esters, ethers, aldehydes, *etc.*
- (c) haloalkyl groups, wherein the halide can be later displaced with a nucleophilic group such as, for example, an amine, a carboxylate anion, thiol anion, carbanion, or an alkoxide ion, thereby resulting in the covalent attachment of a new group at the functional group of the halogen atom;
- (d) dienophile groups, which are capable of participating in Diels-Alder reactions such as, for example, maleimido groups;
- (e) aldehyde or ketone groups, such that subsequent derivatization is possible via formation of carbonyl derivatives such as, for example, imines, hydrazones,

semicarbazones or oximes, or via such mechanisms as Grignard addition or alkyllithium addition;

(f) sulfonyl halide groups for subsequent reaction with amines, for example, to form sulfonamides;

5 (g) thiol groups, which can be, for example, converted to disulfides or reacted with acyl halides;

(h) amine or sulfhydryl groups, which can be, for example, acylated, alkylated or oxidized;

10 (i) alkenes, which can undergo, for example, cycloadditions, acylation, Michael addition, *etc*; and

(j) epoxides, which can react with, for example, amines and hydroxyl compounds.

[0266] The reactive functional groups can be chosen such that they do not participate in, or interfere with, the reactions necessary to assemble the reactive sugar nucleus or modifying group. Alternatively, a reactive functional group can be protected from participating in the reaction by the presence of a protecting group. Those of skill in the art understand how to protect a particular functional group such that it does not interfere with a chosen set of reaction conditions. For examples of useful protecting groups, *see*, for example, Greene *et al.*, PROTECTIVE GROUPS IN ORGANIC SYNTHESIS, John Wiley & Sons, New York, 1991.

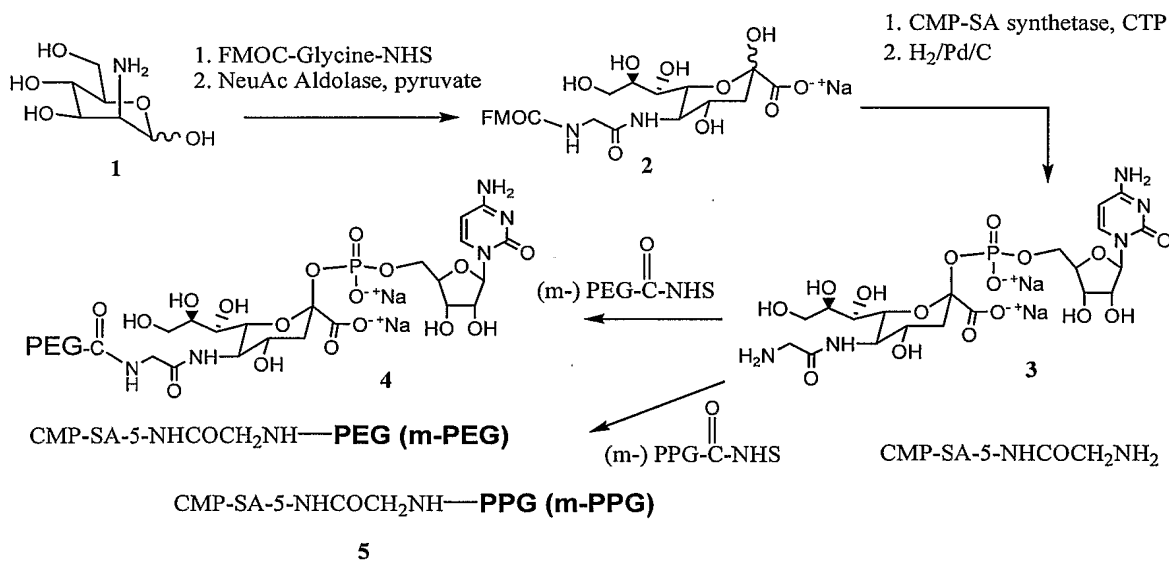
20 [0267] In the discussion that follows, a number of specific examples of modified sugars that are useful in practicing the present invention are set forth. In the exemplary embodiments, a sialic acid derivative is utilized as the sugar nucleus to which the modifying group is attached. The focus of the discussion on sialic acid derivatives is for clarity of illustration only and should not be construed to limit the scope of the invention. Those of skill in the art will appreciate that a variety of other sugar moieties can be activated and derivatized in a manner analogous to that set forth using sialic acid as an example. For example, numerous methods are available for modifying galactose, glucose, N-acetylgalactosamine and fucose to name a few sugar substrates, which are readily modified by art recognized methods. *See*, for example, Elhalabi *et al.*, *Curr. Med. Chem.* **6**: 93 (1999); and Schafer *et al.*, *J. Org. Chem.* **65**: 24 (2000)).

30 [0268] In an exemplary embodiment, the peptide that is modified by a method of the invention is a GLP-1 peptide that has had one or more mutations introduced according to the

methods of the invention. The oligosaccharide chains of the glycopeptide lacking a sialic acid and containing a terminal galactose residue can be glyco-PEG-ylated, glyco-PPG-ylated or otherwise modified with a modified sialic acid.

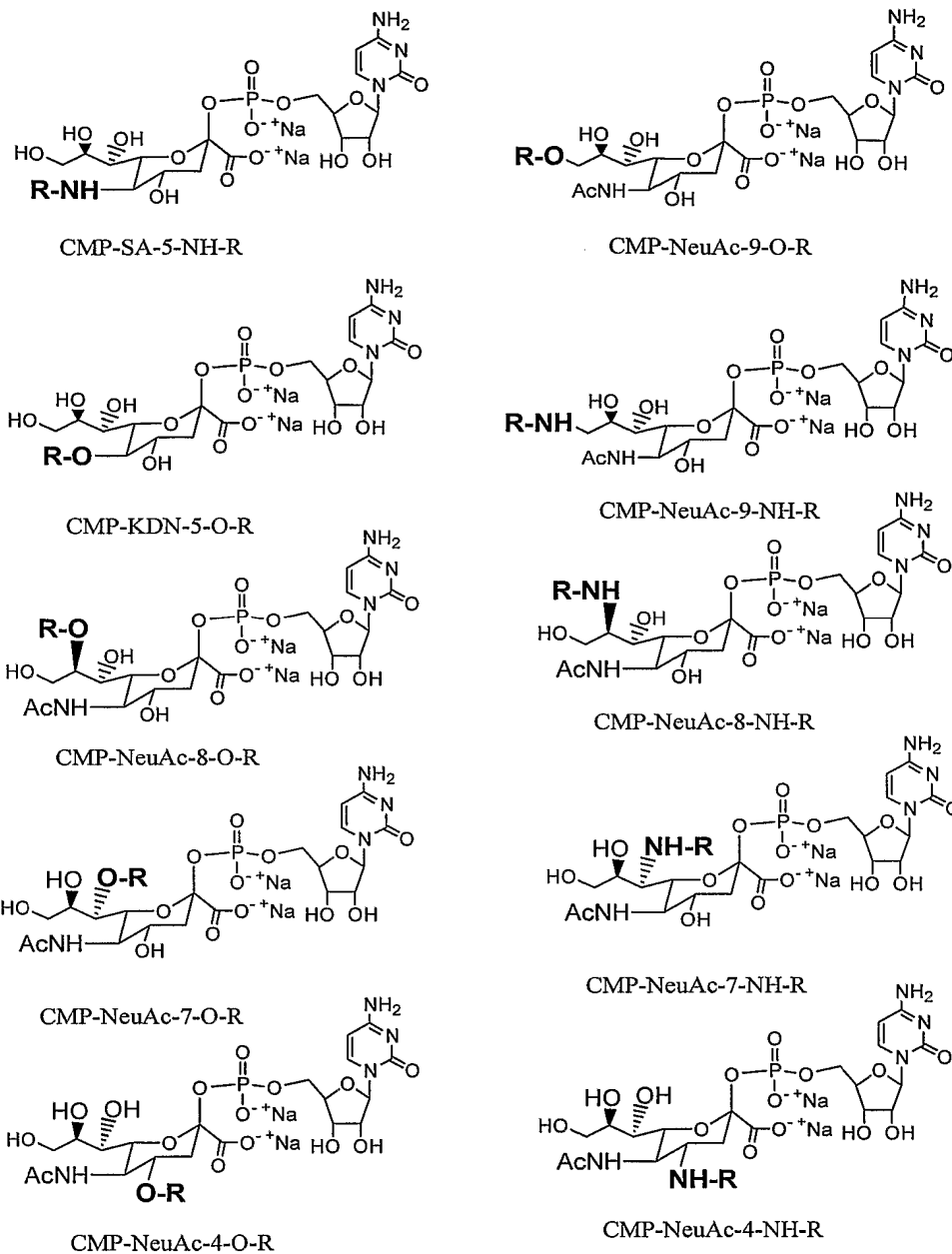
[0269] In Scheme 1, the amino glycoside **1**, is treated with the active ester of a protected amino acid (*e.g.*, glycine) derivative, converting the sugar amine residue into the corresponding protected amino acid amide adduct. The adduct is treated with an aldolase to form α -hydroxy carboxylate **2**. Compound **2** is converted to the corresponding CMP derivative by the action of CMP-SA synthetase, followed by catalytic hydrogenation of the CMP derivative to produce compound **3**. The amine introduced via formation of the glycine adduct is utilized as a locus of PEG or PPG attachment by reacting compound **3** with an activated (*m*-) PEG or (*m*-) PPG derivative (*e.g.*, PEG-C(O)NHS, PPG-C(O)NHS), producing **4** or **5**, respectively.

Scheme 1

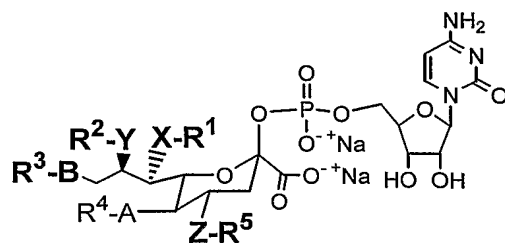


[0270] Table 1 sets forth representative examples of sugar monophosphates that are derivatized with a PEG or PPG moiety. Certain of the compounds of Table 1 are prepared by the method of Scheme 1. Other derivatives are prepared by art-recognized methods. *See*, for example, Keppler *et al.*, *Glycobiology* **11**: 11R (2001); and Charter *et al.*, *Glycobiology* **10**: 1049 (2000)). Other amine reactive PEG and PPG analogues are commercially available, or they can be prepared by methods readily accessible to those of skill in the art.

Table 1



- 5 [0271] The modified sugar phosphates of use in practicing the present invention can be substituted in other positions as well as those set forth above. Presently preferred substitutions of sialic acid are set forth in Formula I:



(I)

in which X is a linking group, which is preferably selected from -O-, -N(H)-, -S, CH₂-, and -N(R)₂, in which each R is a member independently selected from R¹-R⁵. The symbols Y, Z, A and B each represent a group that is selected from the group set forth above for the identity of X. X, Y, Z, A and B are each independently selected and, therefore, they can be the same or different. The symbols R¹, R², R³, R⁴ and R⁵ represent H, a water-soluble polymer, therapeutic moiety, biomolecule or other moiety. Alternatively, these symbols represent a linker that is bound to a water-soluble polymer, therapeutic moiety, biomolecule or other moiety.

[0272] Exemplary moieties attached to the conjugates disclosed herein include, but are not limited to, PEG derivatives (*e.g.*, alkyl-PEG, acyl-PEG, acyl-alkyl-PEG, alkyl-acyl-PEG carbamoyl-PEG, aryl-PEG), PPG derivatives (*e.g.*, alkyl-PPG, acyl-PPG, acyl-alkyl-PPG, alkyl-acyl-PPG carbamoyl-PPG, aryl-PPG), therapeutic moieties, diagnostic moieties, mannose-6-phosphate, heparin, heparan, SLe_x, mannose, mannose-6-phosphate, Sialyl Lewis X, FGF, VFGF, proteins, chondroitin, keratan, dermatan, albumin, integrins, antennary oligosaccharides, peptides and the like. Methods of conjugating the various modifying groups to a saccharide moiety are readily accessible to those of skill in the art (POLY (ETHYLENE GLYCOL CHEMISTRY : BIOTECHNICAL AND BIOMEDICAL APPLICATIONS, J. Milton Harris, Ed., Plenum Pub. Corp., 1992; POLY (ETHYLENE GLYCOL) CHEMICAL AND BIOLOGICAL APPLICATIONS, J. Milton Harris, Ed., ACS Symposium Series No. 680, American Chemical Society, 1997; Hermanson, BIOCONJUGATE TECHNIQUES, Academic Press, San Diego, 1996; and Dunn *et al.*, Eds. POLYMERIC DRUGS AND DRUG DELIVERY SYSTEMS, ACS Symposium Series Vol. 469, American Chemical Society, Washington, D.C. 1991).

Cross-linking Groups

[0273] Preparation of the modified sugar for use in the methods of the present invention includes attachment of a modifying group to a sugar residue and forming a stable adduct,

which is a substrate for a glycosyltransferase. The sugar and modifying group can be coupled by a zero- or higher-order cross-linking agent. Exemplary bifunctional compounds which can be used for attaching modifying groups to carbohydrate moieties include, but are not limited to, bifunctional poly(ethyleneglycols), polyamides, polyethers, polyesters and the like. General approaches for linking carbohydrates to other molecules are known in the literature. See, for example, Lee *et al.*, *Biochemistry* **28**: 1856 (1989); Bhatia *et al.*, *Anal. Biochem.* **178**: 408 (1989); Janda *et al.*, *J. Am. Chem. Soc.* **112**: 8886 (1990) and Bednarski *et al.*, WO 92/18135. In the discussion that follows, the reactive groups are treated as benign on the sugar moiety of the nascent modified sugar. The focus of the discussion is for clarity of illustration. Those of skill in the art will appreciate that the discussion is relevant to reactive groups on the modifying group as well.

[0274] An exemplary strategy involves incorporation of a protected sulfhydryl onto the sugar using the heterobifunctional crosslinker SPDP (n-succinimidyl-3-(2-pyridyldithio)propionate and then deprotecting the sulfhydryl for formation of a disulfide bond with another sulfhydryl on the modifying group.

[0275] If SPDP detrimentally affects the ability of the modified sugar to act as a glycosyltransferase substrate, one of an array of other crosslinkers such as 2-iminothiolane or N-succinimidyl S-acetylthioacetate (SATA) is used to form a disulfide bond. 2-iminothiolane reacts with primary amines, instantly incorporating an unprotected sulfhydryl onto the amine-containing molecule. SATA also reacts with primary amines, but incorporates a protected sulfhydryl, which is later deacetylated using hydroxylamine to produce a free sulfhydryl. In each case, the incorporated sulfhydryl is free to react with other sulfhydryls or protected sulfhydryl, like SPDP, forming the required disulfide bond.

[0276] The above-described strategy is exemplary, and not limiting, of linkers of use in the invention. Other crosslinkers are available that can be used in different strategies for crosslinking the modifying group to the peptide. For example, TPCH(S-(2-thiopyridyl)-L-cysteine hydrazide and TPMPH ((S-(2-thiopyridyl) mercapto-propionohydrazide) react with carbohydrate moieties that have been previously oxidized by mild periodate treatment, thus forming a hydrazone bond between the hydrazide portion of the crosslinker and the periodate generated aldehydes. TPCH and TPMPH introduce a 2-pyridylthione protected sulfhydryl group onto the sugar, which can be deprotected with DTT and then subsequently used for conjugation, such as forming disulfide bonds between components.

[0277] If disulfide bonding is found unsuitable for producing stable modified sugars, other crosslinkers may be used that incorporate more stable bonds between components. The heterobifunctional crosslinkers GMBS (N-gamma-malimidobutyryloxy)succinimide) and SMCC (succinimidyl 4-(N-maleimido-methyl)cyclohexane) react with primary amines, thus
5 introducing a maleimide group onto the component. The maleimide group can subsequently react with sulfhydryls on the other component, which can be introduced by previously mentioned crosslinkers, thus forming a stable thioether bond between the components. If steric hindrance between components interferes with either component's activity or the ability of the modified sugar to act as a glycosyltransferase substrate, crosslinkers can be used which
10 introduce long spacer arms between components and include derivatives of some of the previously mentioned crosslinkers (*i.e.*, SPDP). Thus, there is an abundance of suitable crosslinkers, which are useful; each of which is selected depending on the effects it has on optimal peptide conjugate and modified sugar production.

[0278] A variety of reagents are used to modify the components of the modified sugar with
15 intramolecular chemical crosslinks (for reviews of crosslinking reagents and crosslinking procedures see: Wold, F., *Meth. Enzymol.* **25**: 623-651, 1972; Weetall, H. H., and Cooney, D. A., In: *ENZYMES AS DRUGS*. (Holcenberg, and Roberts, eds.) pp. 395-442, Wiley, New York, 1981; Ji, T. H., *Meth. Enzymol.* **91**: 580-609, 1983; Mattson *et al.*, *Mol. Biol. Rep.* **17**: 167-183, 1993, all of which are incorporated herein by reference). Preferred crosslinking reagents
20 are derived from various zero-length, homo-bifunctional, and hetero-bifunctional crosslinking reagents. Zero-length crosslinking reagents include direct conjugation of two intrinsic chemical groups with no introduction of extrinsic material. Agents that catalyze formation of a disulfide bond belong to this category. Another example is reagents that induce condensation of a carboxyl and a primary amino group to form an amide bond such as
25 carbodiimides, ethylchloroformate, Woodward's reagent K (2-ethyl-5-phenylisoxazolium-3'-sulfonate), and carbonyldiimidazole. In addition to these chemical reagents, the enzyme transglutaminase (glutamyl-peptide γ -glutamyltransferase; EC 2.3.2.13) may be used as zero-length crosslinking reagent. This enzyme catalyzes acyl transfer reactions at carboxamide groups of protein-bound glutamyl residues, usually with a primary amino group as
30 substrate. Preferred homo- and hetero-bifunctional reagents contain two identical or two dissimilar sites, respectively, which may be reactive for amino, sulfhydryl, guanidino, indole, or nonspecific groups.

*i. Preferred Specific Sites in Crosslinking Reagents**1. Amino-Reactive Groups*

[0279] In one preferred embodiment, the sites on the cross-linker are amino-reactive groups. Useful non-limiting examples of amino-reactive groups include N-

5 hydroxysuccinimide (NHS) esters, imidoesters, isocyanates, acylhalides, arylazides, p-nitrophenyl esters, aldehydes, and sulfonyl chlorides.

[0280] NHS esters react preferentially with the primary (including aromatic) amino groups of a modified sugar component. The imidazole groups of histidines are known to compete with primary amines for reaction, but the reaction products are unstable and readily
10 hydrolyzed. The reaction involves the nucleophilic attack of an amine on the acid carboxyl of an NHS ester to form an amide, releasing the N-hydroxysuccinimide. Thus, the positive charge of the original amino group is lost.

[0281] Imidoesters are the most specific acylating reagents for reaction with the amine groups of the modified sugar components. At a pH between 7 and 10, imidoesters react only
15 with primary amines. Primary amines attack imidates nucleophilically to produce an intermediate that breaks down to amidine at high pH or to a new imidate at low pH. The new imidate can react with another primary amine, thus crosslinking two amino groups, a case of a putatively monofunctional imidate reacting bifunctionally. The principal product of reaction with primary amines is an amidine that is a stronger base than the original amine.
20 The positive charge of the original amino group is therefore retained.

[0282] Isocyanates (and isothiocyanates) react with the primary amines of the modified sugar components to form stable bonds. Their reactions with sulfhydryl, imidazole, and tyrosyl groups give relatively unstable products.

[0283] Acylazides are also used as amino-specific reagents in which nucleophilic amines of
25 the affinity component attack acidic carboxyl groups under slightly alkaline conditions, *e.g.* pH 8.5.

[0284] Arylhalides such as 1,5-difluoro-2,4-dinitrobenzene react preferentially with the amino groups and tyrosine phenolic groups of modified sugar components, but also with sulfhydryl and imidazole groups.

[0285] p-Nitrophenyl esters of mono- and dicarboxylic acids are also useful amino-reactive groups. Although the reagent specificity is not very high, α - and ϵ -amino groups appear to react most rapidly.

[0286] Aldehydes such as glutaraldehyde react with primary amines of modified sugar.

5 Although unstable Schiff bases are formed upon reaction of the amino groups with the aldehydes of the aldehydes, glutaraldehyde is capable of modifying the modified sugar with stable crosslinks. At pH 6-8, the pH of typical crosslinking conditions, the cyclic polymers undergo a dehydration to form α - β unsaturated aldehyde polymers. Schiff bases, however, are stable, when conjugated to another double bond. The resonant interaction of both double
10 bonds prevents hydrolysis of the Schiff linkage. Furthermore, amines at high local concentrations can attack the ethylenic double bond to form a stable Michael addition product.

[0287] Aromatic sulfonyl chlorides react with a variety of sites of the modified sugar components, but reaction with the amino groups is the most important, resulting in a stable
15 sulfonamide linkage.

2. Sulfhydryl-Reactive Groups

[0288] In another preferred embodiment, the sites are sulfhydryl-reactive groups. Useful, non-limiting examples of sulfhydryl-reactive groups include maleimides, alkyl halides, pyridyl disulfides, and thiophthalimides.

20 [0289] Maleimides react preferentially with the sulfhydryl group of the modified sugar components to form stable thioether bonds. They also react at a much slower rate with primary amino groups and the imidazole groups of histidines. However, at pH 7 the maleimide group can be considered a sulfhydryl-specific group, since at this pH the reaction rate of simple thiols is 1000-fold greater than that of the corresponding amine.

25 [0290] Alkyl halides react with sulfhydryl groups, sulfides, imidazoles, and amino groups. At neutral to slightly alkaline pH, however, alkyl halides react primarily with sulfhydryl groups to form stable thioether bonds. At higher pH, reaction with amino groups is favored.

[0291] Pyridyl disulfides react with free sulfhydryls via disulfide exchange to give mixed disulfides. As a result, pyridyl disulfides are the most specific sulfhydryl-reactive groups.

30 [0292] Thiophthalimides react with free sulfhydryl groups to form disulfides.

3. Carboxyl-Reactive Residue

[0293] In another embodiment, carbodiimides soluble in both water and organic solvent, are used as carboxyl-reactive reagents. These compounds react with free carboxyl groups forming a pseudourea that can then couple to available amines yielding an amide linkage
5 teach how to modify a carboxyl group with carbodiimide (Yamada *et al.*, *Biochemistry* **20**: 4836-4842, 1981).

ii. Preferred Nonspecific Sites in Crosslinking Reagents

[0294] In addition to the use of site-specific reactive moieties, the present invention contemplates the use of non-specific reactive groups to link the sugar to the modifying group.

10 [0295] Exemplary non-specific cross-linkers include photoactivatable groups, completely inert in the dark, which are converted to reactive species upon absorption of a photon of appropriate energy. In one preferred embodiment, photoactivatable groups are selected from precursors of nitrenes generated upon heating or photolysis of azides. Electron-deficient nitrenes are extremely reactive and can react with a variety of chemical bonds including N-H,
15 O-H, C-H, and C=C. Although three types of azides (aryl, alkyl, and acyl derivatives) may be employed, arylazides are presently preferred. The reactivity of arylazides upon photolysis is better with N-H and O-H than C-H bonds. Electron-deficient arylnitrenes rapidly ring-expand to form dehydroazepines, which tend to react with nucleophiles, rather than form C-H insertion products. The reactivity of arylazides can be increased by the presence of electron-
20 withdrawing substituents such as nitro or hydroxyl groups in the ring. Such substituents push the absorption maximum of arylazides to longer wavelength. Unsubstituted arylazides have an absorption maximum in the range of 260-280 nm, while hydroxy and nitroarylazides absorb significant light beyond 305 nm. Therefore, hydroxy and nitroarylazides are most preferable since they allow to employ less harmful photolysis conditions for the affinity
25 component than unsubstituted arylazides.

[0296] In another preferred embodiment, photoactivatable groups are selected from fluorinated arylazides. The photolysis products of fluorinated arylazides are arylnitrenes, all of which undergo the characteristic reactions of this group, including C-H bond insertion, with high efficiency (Keana *et al.*, *J. Org. Chem.* **55**: 3640-3647, 1990).

30 [0297] In another embodiment, photoactivatable groups are selected from benzophenone residues. Benzophenone reagents generally give higher crosslinking yields than arylazide reagents.

[0298] In another embodiment, photoactivatable groups are selected from diazo compounds, which form an electron-deficient carbene upon photolysis. These carbenes undergo a variety of reactions including insertion into C-H bonds, addition to double bonds (including aromatic systems), hydrogen abstraction and coordination to nucleophilic centers to give carbon ions.

[0299] In still another embodiment, photoactivatable groups are selected from diazopyruvates. For example, the p-nitrophenyl ester of p-nitrophenyl diazopyruvate reacts with aliphatic amines to give diazopyruvic acid amides that undergo ultraviolet photolysis to form aldehydes. The photolyzed diazopyruvate-modified affinity component will react like formaldehyde or glutaraldehyde forming crosslinks.

iii. Homobifunctional Reagents

1. Homobifunctional crosslinkers reactive with primary amines

[0300] Synthesis, properties, and applications of amine-reactive cross-linkers are commercially described in the literature (for reviews of crosslinking procedures and reagents, *see above*). Many reagents are available (*e.g.*, Pierce Chemical Company, Rockford, Ill.; Sigma Chemical Company, St. Louis, Mo.; Molecular Probes, Inc., Eugene, OR.).

[0301] Preferred, non-limiting examples of homobifunctional NHS esters include disuccinimidyl glutarate (DSG), disuccinimidyl suberate (DSS), bis(sulfosuccinimidyl) suberate (BS), disuccinimidyl tartarate (DST), disulfosuccinimidyl tartarate (sulfo-DST), bis-2-(succinimidooxycarbonyloxy)ethylsulfone (BSOCOES), bis-2-(sulfosuccinimidooxycarbonyloxy)ethylsulfone (sulfo-BSOCOES), ethylene glycolbis(succinimidylsuccinate) (EGS), ethylene glycolbis(sulfosuccinimidylsuccinate) (sulfo-EGS), dithiobis(succinimidylpropionate) (DSP), and dithiobis(sulfosuccinimidylpropionate) (sulfo-DSP). Preferred, non-limiting examples of homobifunctional imidoesters include dimethyl malonimide (DMM), dimethyl succinimide (DMSC), dimethyl adipimide (DMA), dimethyl pimelimide (DMP), dimethyl suberimide (DMS), dimethyl-3,3'-oxydipropionimide (DODP), dimethyl-3,3'-(methylenedioxy)dipropionimide (DMDP), dimethyl-3,3'-(dimethylenedioxy)dipropionimide (DDDP), dimethyl-3,3'-(tetramethylenedioxy)-dipropionimide (DTDP), and dimethyl-3,3'-dithiobispropionimide (DTBP).

[0302] Preferred, non-limiting examples of homobifunctional isothiocyanates include: p-phenylenediisothiocyanate (DITC), and 4,4'-diisothiocyano-2,2'-disulfonic acid stilbene (DIDS).

[0303] Preferred, non-limiting examples of homobifunctional isocyanates include xylene-diisocyanate, toluene-2,4-diisocyanate, toluene-2-isocyanate-4-isothiocyanate, 3-methoxydiphenylmethane-4,4'-diisocyanate, 2,2'-dicarboxy-4,4'-azophenyldiisocyanate, and hexamethylenediisocyanate.

5 [0304] Preferred, non-limiting examples of homobifunctional arylhalides include 1,5-difluoro-2,4-dinitrobenzene (DFDNB), and 4,4'-difluoro-3,3'-dinitrophenyl-sulfone.

[0305] Preferred, non-limiting examples of homobifunctional aliphatic aldehyde reagents include glyoxal, malondialdehyde, and glutaraldehyde.

10 [0306] Preferred, non-limiting examples of homobifunctional acylating reagents include nitrophenyl esters of dicarboxylic acids.

[0307] Preferred, non-limiting examples of homobifunctional aromatic sulfonyl chlorides include phenol-2,4-disulfonyl chloride, and α -naphthol-2,4-disulfonyl chloride.

[0308] Preferred, non-limiting examples of additional amino-reactive homobifunctional reagents include erythritolbiscarbonate which reacts with amines to give biscarbamates.

15 2. Homobifunctional Crosslinkers Reactive with Free Sulfhydryl Groups

[0309] Synthesis, properties, and applications of such reagents are described in the literature (for reviews of crosslinking procedures and reagents, *see above*). Many of the reagents are commercially available (*e.g.*, Pierce Chemical Company, Rockford, Ill.; Sigma Chemical Company, St. Louis, Mo.; Molecular Probes, Inc., Eugene, OR).

20 [0310] Preferred, non-limiting examples of homobifunctional maleimides include bismaleimido-hexane (BMH), N,N'-(1,3-phenylene) bismaleimide, N,N'-(1,2-phenylene)bismaleimide, azophenyldimaleimide, and bis(N-maleimidomethyl)ether.

[0311] Preferred, non-limiting examples of homobifunctional pyridyl disulfides include 1,4-di-3'-(2'-pyridyldithio)propionamidobutane (DPDPB).

25 [0312] Preferred, non-limiting examples of homobifunctional alkyl halides include 2,2'-dicarboxy-4,4'-diiodoacetamidoazobenzene, α,α' -diiodo-p-xylenesulfonic acid, α,α' -dibromo-p-xylenesulfonic acid, N,N'-bis(b-bromoethyl)benzylamine, N,N'-di(bromoacetyl)phenylthydrazine, and 1,2-di(bromoacetyl)amino-3-phenylpropane.

3. Homobifunctional Photoactivatable Crosslinkers

[0313] Synthesis, properties, and applications of such reagents are described in the literature (for reviews of crosslinking procedures and reagents, *see above*). Some of the reagents are commercially available (*e.g.*, Pierce Chemical Company, Rockford, Ill.; Sigma Chemical Company, St. Louis, Mo.; Molecular Probes, Inc., Eugene, OR).

[0314] Preferred, non-limiting examples of homobifunctional photoactivatable crosslinker include bis- β -(4-azidosalicylamido)ethyl disulfide (BASED), di-N-(2-nitro-4-azidophenyl)-cystamine-S,S-dioxide (DNCO), and 4,4'-dithiobisphenylazide.

iv. HeteroBifunctional Reagents

1. Amino-Reactive HeteroBifunctional Reagents with a Pyridyl Disulfide Moiety

[0315] Synthesis, properties, and applications of such reagents are described in the literature (for reviews of crosslinking procedures and reagents, *see above*). Many of the reagents are commercially available (*e.g.*, Pierce Chemical Company, Rockford, Ill.; Sigma Chemical Company, St. Louis, Mo.; Molecular Probes, Inc., Eugene, OR).

[0316] Preferred, non-limiting examples of hetero-bifunctional reagents with a pyridyl disulfide moiety and an amino-reactive NHS ester include N-succinimidyl-3-(2-pyridyldithio)propionate (SPDP), succinimidyl 6-3-(2-pyridyldithio)propionamido hexanoate (LC-SPDP), sulfosuccinimidyl 6-3-(2-pyridyldithio)propionamido hexanoate (sulfo-LCSPDP), 4-succinimidyl oxycarbonyl- α -methyl- α -(2-pyridyldithio)toluene (SMPT), and sulfosuccinimidyl 6- α -methyl- α -(2-pyridyldithio)toluamido hexanoate (sulfo-LC-SMPT).

2. Amino-Reactive HeteroBifunctional Reagents with a Maleimide Moiety

[0317] Synthesis, properties, and applications of such reagents are described in the literature. Preferred, non-limiting examples of hetero-bifunctional reagents with a maleimide moiety and an amino-reactive NHS ester include succinimidyl maleimidylacetate (AMAS), succinimidyl 3-maleimidylpropionate (BMPS), N- γ -maleimidobutyryloxysuccinimide ester (GMBS) N- γ -maleimidobutyryloxysulfo succinimide ester (sulfo-GMBS) succinimidyl 6-maleimidylhexanoate (EMCS), succinimidyl 3-maleimidylbenzoate (SMB), m-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS), m-maleimidobenzoyl-N-hydroxysulfosuccinimide ester (sulfo-MBS), succinimidyl 4-(N-maleimido methyl)-cyclohexane-1-carboxylate (SMCC), sulfosuccinimidyl 4-(N-maleimido methyl)cyclohexane-

1-carboxylate (sulfo-SMCC), succinimidyl 4-(p-maleimidophenyl)butyrate (SMPB), and sulfosuccinimidyl 4-(p-maleimidophenyl)butyrate (sulfo-SMPB).

3. Amino-Reactive HeteroBifunctional Reagents with an Alkyl Halide Moiety

[0318] Synthesis, properties, and applications of such reagents are described in the literature Preferred, non-limiting examples of hetero-bifunctional reagents with an alkyl halide moiety and an amino-reactive NHS ester include N-succinimidyl-(4-iodoacetyl)aminobenzoate (SIAB), sulfosuccinimidyl-(4-iodoacetyl)aminobenzoate (sulfo-SIAB), succinimidyl-6-(iodoacetyl)aminohexanoate (SIAX), succinimidyl-6-(6-((iodoacetyl)-amino)hexanoylamino)hexanoate (SIAXX), succinimidyl-6-(((4-(iodoacetyl)-amino)-methyl)-cyclohexane-1-carbonyl)aminohexanoate (SIACX), and succinimidyl-4((iodoacetyl)-amino)methylcyclohexane-1-carboxylate (SIAC).

[0319] A preferred example of a hetero-bifunctional reagent with an amino-reactive NHS ester and an alkyl dihalide moiety is N-hydroxysuccinimidyl 2,3-dibromopropionate (SDBP). SDBP introduces intramolecular crosslinks to the affinity component by conjugating its amino groups. The reactivity of the dibromopropionyl moiety towards primary amine groups is controlled by the reaction temperature (McKenzie *et al.*, *Protein Chem.* **7**: 581-592 (1988)).

[0320] Preferred, non-limiting examples of hetero-bifunctional reagents with an alkyl halide moiety and an amino-reactive p-nitrophenyl ester moiety include p-nitrophenyl iodoacetate (NPIA).

[0321] Other cross-linking agents are known to those of skill in the art. *See*, for example, Pomato *et al.*, U.S. Patent No. 5,965,106. It is within the abilities of one of skill in the art to choose an appropriate cross-linking agent for a particular application.

v. Cleavable Linker Groups

[0322] In yet a further embodiment, the linker group is provided with a group that can be cleaved to release the modifying group from the sugar residue. Many cleaveable groups are known in the art. *See*, for example, Jung *et al.*, *Biochem. Biophys. Acta* **761**: 152-162 (1983); Joshi *et al.*, *J. Biol. Chem.* **265**: 14518-14525 (1990); Zarling *et al.*, *J. Immunol.* **124**: 913-920 (1980); Bouiziar *et al.*, *Eur. J. Biochem.* **155**: 141-147 (1986); Park *et al.*, *J. Biol. Chem.* **261**: 205-210 (1986); Browning *et al.*, *J. Immunol.* **143**: 1859-1867 (1989). Moreover a broad

range of cleavable, bifunctional (both homo- and hetero-bifunctional) linker groups is commercially available from suppliers such as Pierce.

[0323] Exemplary cleaveable moieties can be cleaved using light, heat or reagents such as thiols, hydroxylamine, bases, periodate and the like. Moreover, certain preferred groups are cleaved *in vivo* in response to being endocytized (*e.g.*, cis-aconityl; *see*, Shen *et al.*, *Biochem. Biophys. Res. Commun.* **102**: 1048 (1991)). Preferred cleaveable groups comprise a cleaveable moiety which is a member selected from the group consisting of disulfide, ester, imide, carbonate, nitrobenzyl, phenacyl and benzoin groups.

Conjugation of Modified Sugars to Peptides

[0324] The modified sugars are conjugated to a glycosylated or non-glycosylated peptide using an appropriate enzyme to mediate the conjugation. Preferably, the concentrations of the modified donor sugar(s), enzyme(s) and acceptor peptide(s) are selected such that glycosylation proceeds until the acceptor is consumed. The considerations discussed below, while set forth in the context of a sialyltransferase, are generally applicable to other glycosyltransferase reactions.

[0325] A number of methods of using glycosyltransferases to synthesize desired oligosaccharide structures are known and are generally applicable to the instant invention. Exemplary methods are described, for instance, WO 96/32491, Ito *et al.*, *Pure Appl. Chem.* **65**: 753 (1993), and U.S. Pat. Nos. 5,352,670, 5,374,541, and 5,545,553.

[0326] The present invention is practiced using a single glycosyltransferase or a combination of glycosyltransferases. For example, one can use a combination of a sialyltransferase and a galactosyltransferase. In those embodiments using more than one enzyme, the enzymes and substrates are preferably combined in an initial reaction mixture, or the enzymes and reagents for a second enzymatic reaction are added to the reaction medium once the first enzymatic reaction is complete or nearly complete. By conducting two enzymatic reactions in sequence in a single vessel, overall yields are improved over procedures in which an intermediate species is isolated. Moreover, cleanup and disposal of extra solvents and by-products is reduced.

[0327] In a preferred embodiment, each of the first and second enzyme is a glycosyltransferase. In another preferred embodiment, one enzyme is an endoglycosidase. In an additional preferred embodiment, more than two enzymes are used to assemble the modified glycoprotein of the invention. The enzymes are used to alter a saccharide structure

on the peptide at any point either before or after the addition of the modified sugar to the peptide.

[0328] The O-linked glycosyl moieties of the conjugates of the invention are generally originate with a GalNAc moiety that is attached to the peptide. Any member of the family of GalNAc transferases can be used to bind a GalNAc moiety to the peptide (Hassan H, Bennett EP, Mandel U, Hollingsworth MA, and Clausen H (2000). Control of Mucin-Type O-Glycosylation: O-Glycan Occupancy is Directed by Substrate Specificities of Polypeptide GalNAc-Transferases. (Eds. Ernst, Hart, and Sinay). Wiley-VCH chapter "Carbohydrates in Chemistry and Biology - a Comprehensive Handbook", 273-292). The GalNAc moiety itself can be the intact glycosyl linker. Alternatively, the saccharyl residue is built out using one more enzyme and one or more appropriate glycosyl substrate for the enzyme, the modified sugar being added to the built out glycosyl moiety.

[0329] In another embodiment, the method makes use of one or more exo- or endoglycosidase. The glycosidase is typically a mutant, which is engineered to form glycosyl bonds rather than cleave them. The mutant glycanase typically includes a substitution of an amino acid residue for an active site acidic amino acid residue. For example, when the endoglycanase is endo-H, the substituted active site residues will typically be Asp at position 130, Glu at position 132 or a combination thereof. The amino acids are generally replaced with serine, alanine, asparagine, or glutamine.

[0330] The mutant enzyme catalyzes the reaction, usually by a synthesis step that is analogous to the reverse reaction of the endoglycanase hydrolysis step. In these embodiments, the glycosyl donor molecule (*e.g.*, a desired oligo- or mono-saccharide structure) contains a leaving group and the reaction proceeds with the addition of the donor molecule to a GlcNAc residue on the protein. For example, the leaving group can be a halogen, such as fluoride. In other embodiments, the leaving group is a Asn, or a Asn-peptide moiety. In yet further embodiments, the GlcNAc residue on the glycosyl donor molecule is modified. For example, the GlcNAc residue may comprise a 1,2 oxazoline moiety.

[0331] In a preferred embodiment, each of the enzymes utilized to produce a conjugate of the invention are present in a catalytic amount. The catalytic amount of a particular enzyme varies according to the concentration of that enzyme's substrate as well as to reaction conditions such as temperature, time and pH value. Means for determining the catalytic

amount for a given enzyme under preselected substrate concentrations and reaction conditions are well known to those of skill in the art.

[0332] The temperature at which an above process is carried out can range from just above freezing to the temperature at which the most sensitive enzyme denatures. Preferred temperature ranges are about 0 °C to about 55 °C, and more preferably about 20 °C to about 30 °C. In another exemplary embodiment, one or more components of the present method are conducted at an elevated temperature using a thermophilic enzyme.

[0333] The reaction mixture is maintained for a period of time sufficient for the acceptor to be glycosylated, thereby forming the desired conjugate. Some of the conjugate can often be detected after a few hours, with recoverable amounts usually being obtained within 24 hours or less. Those of skill in the art understand that the rate of reaction is dependent on a number of variable factors (*e.g.*, enzyme concentration, donor concentration, acceptor concentration, temperature, solvent volume), which are optimized for a selected system.

[0334] The present invention also provides for the industrial-scale production of modified peptides. As used herein, an industrial scale generally produces at least about 250 mg, preferably at least about 500 mg, and more preferably at least about 1 gram of finished, purified conjugate, preferably after a single reaction cycle, *i.e.*, the conjugate is not a combination the reaction products from identical, consecutively iterated synthesis cycles.

[0335] In the discussion that follows, the invention is exemplified by the conjugation of modified sialic acid moieties to a glycosylated peptide. The exemplary modified sialic acid is labeled with (m-) PEG. The focus of the following discussion on the use of PEG-modified sialic acid and glycosylated peptides is for clarity of illustration and is not intended to imply that the invention is limited to the conjugation of these two partners. One of skill understands that the discussion is generally applicable to the additions of modified glycosyl moieties other than sialic acid. Moreover, the discussion is equally applicable to the modification of a glycosyl unit with agents other than PEG including other water-soluble polymers, therapeutic moieties, and biomolecules.

[0336] An enzymatic approach can be used for the selective introduction of (m-) PEG-ylated or (m-) PPG-ylated carbohydrates onto a peptide or glycopeptide. The method utilizes modified sugars containing PEG, PPG, or a masked reactive functional group, and is combined with the appropriate glycosyltransferase or glycosynthase. By selecting the glycosyltransferase that will make the desired carbohydrate linkage and utilizing the modified

sugar as the donor substrate, the PEG or PPG can be introduced directly onto the peptide backbone, onto existing sugar residues of a glycopeptide or onto sugar residues that have been added to a peptide.

[0337] An acceptor for the sialyltransferase is present on the peptide to be modified by the methods of the present invention either as a naturally occurring structure or one placed there recombinantly, enzymatically or chemically. Suitable acceptors, include, for example, galactosyl acceptors such as GalNAc, Gal β 1,4GlcNAc, Gal β 1,4GalNAc, Gal β 1,3GalNAc, lacto-N-tetraose, Gal β 1,3GlcNAc, Gal β 1,3Ara, Gal β 1,6GlcNAc, Gal β 1,4Glc (lactose), and other acceptors known to those of skill in the art (*see, e.g., Paulson et al., J. Biol. Chem.* **253**: 5617-5624 (1978)).

[0338] In one embodiment, an acceptor for the sialyltransferase is present on the glycopeptide to be modified upon *in vivo* synthesis of the glycopeptide. Such glycopeptides can be sialylated using the claimed methods without prior modification of the glycosylation pattern of the glycopeptide. Alternatively, the methods of the invention can be used to sialylate a peptide that does not include a suitable acceptor; one first modifies the peptide to include an acceptor by methods known to those of skill in the art. In an exemplary embodiment, a GalNAc residue is added to an O-linked glycosylation site by the action of a GalNAc transferase. Hassan H, Bennett EP, Mandel U, Hollingsworth MA, and Clausen H (2000). Control of Mucin-Type O-Glycosylation: O-Glycan Occupancy is Directed by Substrate Specificities of Polypeptide GalNAc-Transferases. (Eds. Ernst, Hart, and Sinay). Wiley-VCH chapter "Carbohydrates in Chemistry and Biology - a Comprehension Handbook", 273-292.

[0339] In an exemplary embodiment, the galactosyl acceptor is assembled by attaching a galactose residue to an appropriate acceptor linked to the peptide, *e.g.*, a GalNAc. The method includes incubating the peptide to be modified with a reaction mixture that contains a suitable amount of a galactosyltransferase (*e.g.*, Gal β 1,3 or Gal β 1,4), and a suitable galactosyl donor (*e.g.*, UDP-galactose). The reaction is allowed to proceed substantially to completion or, alternatively, the reaction is terminated when a preselected amount of the galactose residue is added. Other methods of assembling a selected saccharide acceptor will be apparent to those of skill in the art.

[0340] In yet another embodiment, glycopeptide-linked oligosaccharides are first "trimmed," either in whole or in part, to expose either an acceptor for the sialyltransferase or

a moiety to which one or more appropriate residues can be added to obtain a suitable acceptor. Enzymes such as glycosyltransferases and endoglycosidases (*see*, for example U.S. Patent No. 5,716,812) are useful for the attaching and trimming reactions.

[0341] In the discussion that follows, the method of the invention is exemplified by the use of modified sugars having a water-soluble polymer attached thereto. The focus of the discussion is for clarity of illustration. Those of skill will appreciate that the discussion is equally relevant to those embodiments in which the modified sugar bears a therapeutic moiety, biomolecule or the like.

[0342] In an exemplary embodiment, an O-linked carbohydrate residue is “trimmed” prior to the addition of the modified sugar. For example a GalNAc-Gal residue is trimmed back to GalNAc. A modified sugar bearing a water-soluble polymer is conjugated to one or more of the sugar residues exposed by the “trimming.” In one example, a glycopeptide is “trimmed” and a water-soluble polymer is added to the resulting O-side chain amino acid or glycopeptide glycan via a saccharyl moiety, *e.g.*, Sia, Gal or GalNAc moiety conjugated to the water-soluble polymer. The modified saccharyl moiety is attached to an acceptor site on the “trimmed” glycopeptide. Alternatively, an unmodified saccharyl moiety, *e.g.*, Gal can be added the terminus of the O-linked glycan.

[0343] In another exemplary embodiment, a water-soluble polymer is added to a GalNAc residue via a modified sugar having a galactose residue. Alternatively, an unmodified Gal can be added to the terminal GalNAc residue.

[0344] In yet a further example, a water-soluble polymer is added onto a Gal residue using a modified sialic acid.

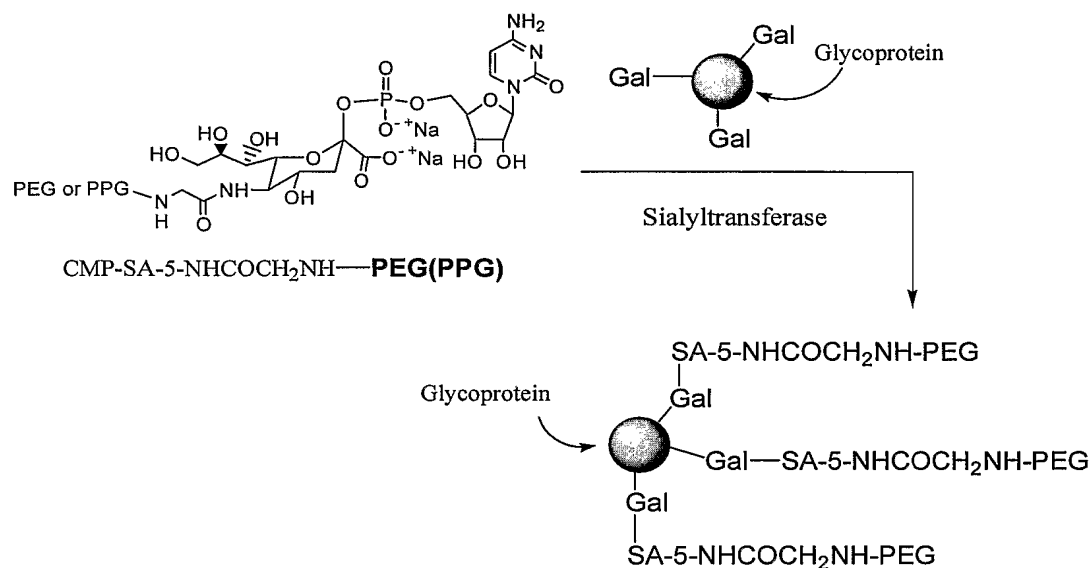
[0345] In another exemplary embodiment, an O-linked glycosyl residue is “trimmed back” to the GalNAc attached to the amino acid. In one example, a water-soluble polymer is added via a Gal modified with the polymer. Alternatively, an unmodified Gal is added to the GalNAc, followed by a Gal with an attached water-soluble polymer. In yet another embodiment, one or more unmodified Gal residue is added to the GalNAc, followed by a sialic acid moiety modified with a water-soluble polymer.

[0346] The exemplary embodiments discussed above provide an illustration of the power of the methods set forth herein. Using the methods of the invention, it is possible to “trim back” and build up a carbohydrate residue of substantially any desired structure. The modified

sugar can be added to the termini of the carbohydrate moiety as set forth above, or it can be intermediate between the peptide core and the terminus of the carbohydrate.

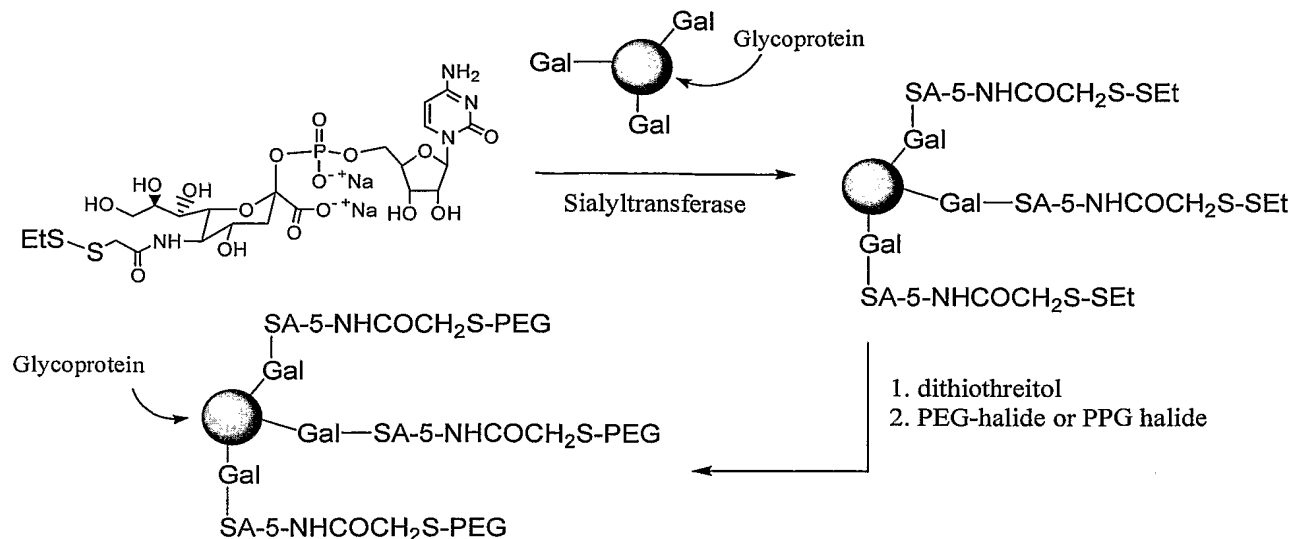
[0347] In an exemplary embodiment, the water-soluble polymer is added to a terminal Gal residue using a polymer modified sialic acid. An appropriate sialyltransferase is used to add a modified sialic acid. The approach is summarized in Scheme 2.

Scheme 2



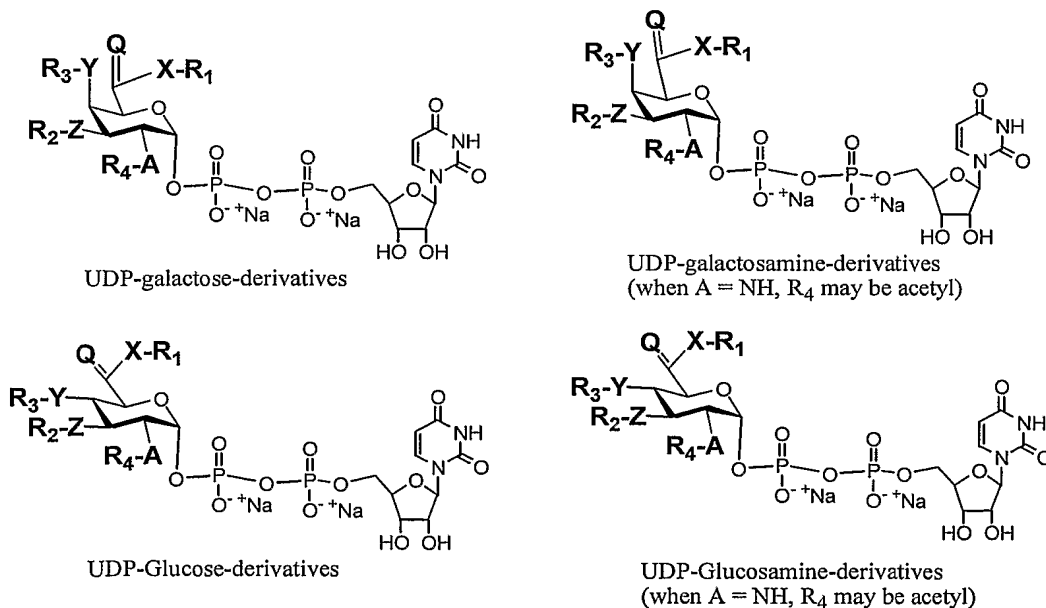
[0348] In yet a further approach, summarized in Scheme 3, a masked reactive functionality is present on the sialic acid. The masked reactive group is preferably unaffected by the conditions used to attach the modified sialic acid to the peptide. After the covalent attachment of the modified sialic acid to the peptide, the mask is removed and the peptide is conjugated with an agent such as PEG, PPG, a therapeutic moiety, biomolecule or other agent. The agent is conjugated to the peptide in a specific manner by its reaction with the unmasked reactive group on the modified sugar residue.

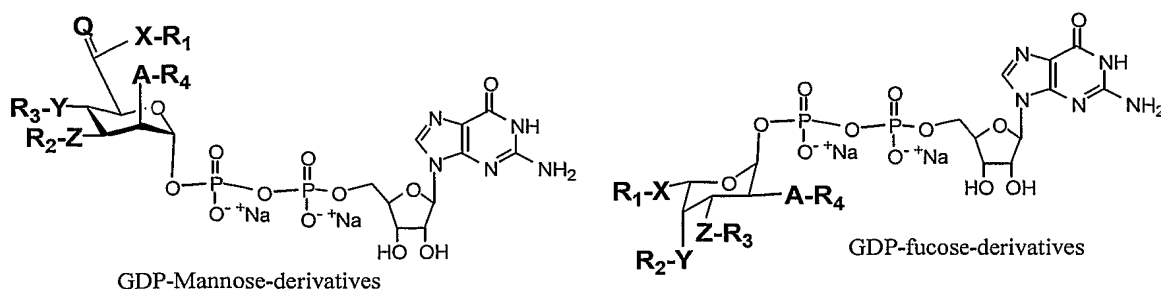
Scheme 3



[0349] Any modified sugar can be used with its appropriate glycosyltransferase, depending on the terminal sugars of the oligosaccharide side chains of the glycopeptide (Table 2). As discussed above, the terminal sugar of the glycopeptide required for introduction of the PEG-ylated or PPGylated structure can be introduced naturally during expression or it can be produced post expression using the appropriate glycosidase(s), glycosyltransferase(s) or mix of glycosidase(s) and glycosyltransferase(s).

Table 2





X = O, NH, S, CH₂, N-(R₁₋₅)₂.
Y = X; Z = X; A = X; B = X.

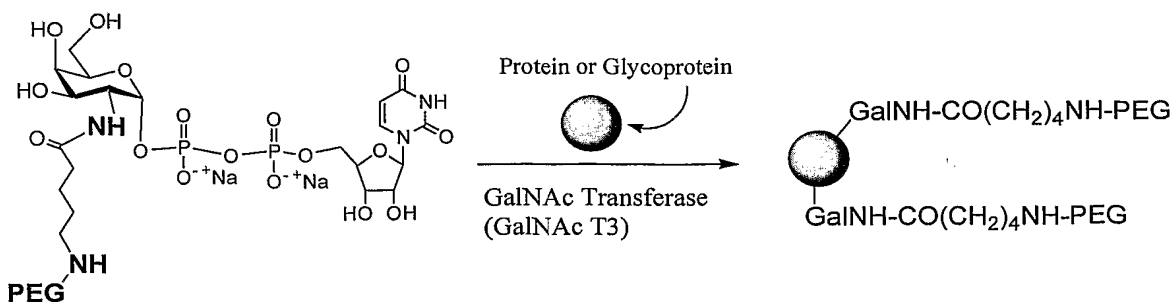
Q = H₂, O, S, NH, N-R.

R, R₁₋₄ = H, Linker-M, M.

M = Ligand of interest

Ligand of interest = acyl-PEG, acyl-PPG, alkyl-PEG, acyl-alkyl-PEG, acyl-alkyl-PEG, carbamoyl-PEG, carbamoyl-PPG, PEG, PPG, acyl-aryl-PEG, acyl-aryl-PPG, aryl-PEG, aryl-PPG, Mannose-6-phosphate, heparin, heparan, SLex, Mannose, FGF, VFGF, protein, chondroitin, keratan, dermatan, albumin, integrins, peptides, etc.

[0350] In an alternative embodiment, the modified sugar is added directly to the peptide backbone using a glycosyltransferase known to transfer sugar residues to the O-linked glycosylation site on the peptide backbone. This exemplary embodiment is set forth in Scheme 4. Exemplary glycosyltransferases useful in practicing the present invention include, but are not limited to, GalNAc transferases (GalNAc T1-20), GlcNAc transferases, fucosyltransferases, glucosyltransferases, xylosyltransferases, mannosyltransferases and the like. Use of this approach allows the direct addition of modified sugars onto peptides that lack any carbohydrates or, alternatively, onto existing glycopeptides. In both cases, the addition of the modified sugar occurs at specific positions on the peptide backbone as defined by the substrate specificity of the glycosyltransferase and not in a random manner as occurs during modification of a protein's peptide backbone using chemical methods. An array of agents can be introduced into proteins or glycopeptides that lack the glycosyltransferase substrate peptide sequence by engineering the appropriate amino acid sequence into the polypeptide chain.

Scheme 4

[0351] In each of the exemplary embodiments set forth above, one or more additional chemical or enzymatic modification steps can be utilized following the conjugation of the modified sugar to the peptide. In an exemplary embodiment, an enzyme (*e.g.*, fucosyltransferase) is used to append a glycosyl unit (*e.g.*, fucose) onto the terminal modified sugar attached to the peptide. In another example, an enzymatic reaction is utilized to “cap” (*e.g.*, sialylate) sites to which the modified sugar failed to conjugate. Alternatively, a chemical reaction is utilized to alter the structure of the conjugated modified sugar. For example, the conjugated modified sugar is reacted with agents that stabilize or destabilize its linkage with the peptide component to which the modified sugar is attached. In another example, a component of the modified sugar is deprotected following its conjugation to the peptide. One of skill will appreciate that there is an array of enzymatic and chemical procedures that are useful in the methods of the invention at a stage after the modified sugar is conjugated to the peptide. Further elaboration of the modified sugar-peptide conjugate is within the scope of the invention.

[0352] In another exemplary embodiment, the glycopeptide is conjugated to a targeting agent, *e.g.*, transferrin (to deliver the peptide across the blood-brain barrier, and to endosomes), carnitine (to deliver the peptide to muscle cells; *see*, for example, LeBorgne *et al.*, *Biochem. Pharmacol.* **59**: 1357-63 (2000), and phosphonates, *e.g.*, bisphosphonate (to target the peptide to bone and other calciferous tissues; *see*, for example, Modern Drug Discovery, August 2002, page 10). Other agents useful for targeting are apparent to those of skill in the art. For example, glucose, glutamine and IGF are also useful to target muscle.

[0353] The targeting moiety and therapeutic peptide are conjugated by any method discussed herein or otherwise known in the art. Those of skill will appreciate that peptides in addition to those set forth above can also be derivatized as set forth herein. Exemplary

peptides are set forth in the Appendix attached to copending, commonly owned US Provisional Patent Application No. 60/328,523 filed October 10, 2001.

[0354] In an exemplary embodiment, the targeting agent and the therapeutic peptide are coupled via a linker moiety. In this embodiment, at least one of the therapeutic peptide or the targeting agent is coupled to the linker moiety via an intact glycosyl linking group according to a method of the invention. In an exemplary embodiment, the linker moiety includes a poly(ether) such as poly(ethylene glycol). In another exemplary embodiment, the linker moiety includes at least one bond that is degraded *in vivo*, releasing the therapeutic peptide from the targeting agent, following delivery of the conjugate to the targeted tissue or region of the body.

[0355] In yet another exemplary embodiment, the *in vivo* distribution of the therapeutic moiety is altered via altering a glycoform on the therapeutic moiety without conjugating the therapeutic peptide to a targeting moiety. For example, the therapeutic peptide can be shunted away from uptake by the reticuloendothelial system by capping a terminal galactose moiety of a glycosyl group with sialic acid (or a derivative thereof).

i. Enzymes

1. Glycosyltransferases

[0356] Glycosyltransferases catalyze the addition of activated sugars (donor NDP-sugars), in a step-wise fashion, to a protein, glycopeptide, lipid or glycolipid or to the non-reducing end of a growing oligosaccharide. N-linked glycopeptides are synthesized via a transferase and a lipid-linked oligosaccharide donor Dol-PP-NAG₂Glc₃Man₉ in an en block transfer followed by trimming of the core. In this case the nature of the "core" saccharide is somewhat different from subsequent attachments. A very large number of glycosyltransferases are known in the art.

[0357] The glycosyltransferase to be used in the present invention may be any as long as it can utilize the modified sugar as a sugar donor. Examples of such enzymes include Leloir pathway glycosyltransferase, such as galactosyltransferase, N-acetylglucosaminyltransferase, N-acetylgalactosaminyltransferase, fucosyltransferase, sialyltransferase, mannosyltransferase, xylosyltransferase, glucurononyltransferase and the like.

[0358] For enzymatic saccharide syntheses that involve glycosyltransferase reactions, glycosyltransferase can be cloned, or isolated from any source. Many cloned glycosyltransferases are known, as are their polynucleotide sequences. *See, e.g.*, "The WWW

Guide To Cloned Glycosyltransferases,” (http://www.vei.co.uk/TGN/gt_guide.htm).

Glycosyltransferase amino acid sequences and nucleotide sequences encoding glycosyltransferases from which the amino acid sequences can be deduced are also found in various publicly available databases, including GenBank, Swiss-Prot, EMBL, and others.

5 [0359] Glycosyltransferases that can be employed in the methods of the invention include, but are not limited to, galactosyltransferases, fucosyltransferases, glucosyltransferases, N-acetylgalactosaminyltransferases, N-acetylglucosaminyltransferases, glucuronyltransferases, sialyltransferases, mannosyltransferases, glucuronic acid transferases, galacturonic acid transferases, and oligosaccharyltransferases. Suitable glycosyltransferases include those
10 obtained from eukaryotes, as well as from prokaryotes.

[0360] DNA encoding glycosyltransferases may be obtained by chemical synthesis, by screening reverse transcripts of mRNA from appropriate cells or cell line cultures, by screening genomic libraries from appropriate cells, or by combinations of these procedures. Screening of mRNA or genomic DNA may be carried out with oligonucleotide probes
15 generated from the glycosyltransferases gene sequence. Probes may be labeled with a detectable group such as a fluorescent group, a radioactive atom or a chemiluminescent group in accordance with known procedures and used in conventional hybridization assays. In the alternative, glycosyltransferases gene sequences may be obtained by use of the polymerase chain reaction (PCR) procedure, with the PCR oligonucleotide primers being produced from
20 the glycosyltransferases gene sequence. *See*, U.S. Pat. No. 4,683,195 to Mullis *et al.* and U.S. Pat. No. 4,683,202 to Mullis.

[0361] The glycosyltransferase may be synthesized in host cells transformed with vectors containing DNA encoding the glycosyltransferases enzyme. Vectors are used either to amplify DNA encoding the glycosyltransferases enzyme and/or to express DNA which
25 encodes the glycosyltransferases enzyme. An expression vector is a replicable DNA construct in which a DNA sequence encoding the glycosyltransferases enzyme is operably linked to suitable control sequences capable of effecting the expression of the glycosyltransferases enzyme in a suitable host. The need for such control sequences will vary depending upon the host selected and the transformation method chosen. Generally,
30 control sequences include a transcriptional promoter, an optional operator sequence to control transcription, a sequence encoding suitable mRNA ribosomal binding sites, and sequences which control the termination of transcription and translation. Amplification vectors do not require expression control domains. All that is needed is the ability to replicate in a host,

usually conferred by an origin of replication, and a selection gene to facilitate recognition of transformants.

[0362] In an exemplary embodiment, the invention utilizes a prokaryotic enzyme. Such glycosyltransferases include enzymes involved in synthesis of lipooligosaccharides (LOS), which are produced by many gram negative bacteria (Preston *et al.*, *Critical Reviews in Microbiology* **23**(3): 139-180 (1996)). Such enzymes include, but are not limited to, the proteins of the *rfa* operons of species such as *E. coli* and *Salmonella typhimurium*, which include a β 1,6 galactosyltransferase and a β 1,3 galactosyltransferase (*see, e.g.*, EMBL Accession Nos. M80599 and M86935 (*E. coli*); EMBL Accession No. S56361 (*S. typhimurium*)), a glucosyltransferase (Swiss-Prot Accession No. P25740 (*E. coli*), an β 1,2-glucosyltransferase (*rfaJ*)(Swiss-Prot Accession No. P27129 (*E. coli*) and Swiss-Prot Accession No. P19817 (*S. typhimurium*)), and an β 1,2-N-acetylglucosaminyltransferase (*rfaK*)(EMBL Accession No. U00039 (*E. coli*). Other glycosyltransferases for which amino acid sequences are known include those that are encoded by operons such as *rfaB*, which have been characterized in organisms such as *Klebsiella pneumoniae*, *E. coli*, *Salmonella typhimurium*, *Salmonella enterica*, *Yersinia enterocolitica*, *Mycobacterium leprosum*, and the *rhI* operon of *Pseudomonas aeruginosa*.

[0363] Also suitable for use in the present invention are glycosyltransferases that are involved in producing structures containing lacto-N-neotetraose, D-galactosyl- β -1,4-N-acetyl-D-glucosaminyl- β -1,3-D-galactosyl- β -1,4-D-glucose, and the P^k blood group trisaccharide sequence, D-galactosyl- α -1,4-D-galactosyl- β -1,4-D-glucose, which have been identified in the LOS of the mucosal pathogens *Neisseria gonorrhoeae* and *N. meningitidis* (Scholten *et al.*, *J. Med. Microbiol.* **41**: 236-243 (1994)). The genes from *N. meningitidis* and *N. gonorrhoeae* that encode the glycosyltransferases involved in the biosynthesis of these structures have been identified from *N. meningitidis* immunotypes L3 and L1 (Jennings *et al.*, *Mol. Microbiol.* **18**: 729-740 (1995)) and the *N. gonorrhoeae* mutant F62 (Gotshlich, *J. Exp. Med.* **180**: 2181-2190 (1994)). In *N. meningitidis*, a locus consisting of three genes, *lgtA*, *lgtB* and *lgE*, encodes the glycosyltransferase enzymes required for addition of the last three of the sugars in the lacto-N-neotetraose chain (Wakarchuk *et al.*, *J. Biol. Chem.* **271**: 19166-73 (1996)). Recently the enzymatic activity of the *lgtB* and *lgtA* gene product was demonstrated, providing the first direct evidence for their proposed glycosyltransferase function (Wakarchuk *et al.*, *J. Biol. Chem.* **271**(45): 28271-276 (1996)). In *N. gonorrhoeae*, there are two additional genes, *lgtD* which adds β -D-GalNAc to the 3 position of the terminal

galactose of the lacto-*N*-neotetraose structure and *lgtC* which adds a terminal α -D-Gal to the lactose element of a truncated LOS, thus creating the P^k blood group antigen structure (Gotshlich (1994), *supra.*). In *N. meningitidis*, a separate immunotype L1 also expresses the P^k blood group antigen and has been shown to carry an *lgtC* gene (Jennings *et al.*, (1995), *supra.*). *Neisseria* glycosyltransferases and associated genes are also described in USPN 5,545,553 (Gotshlich). Genes for α 1,2-fucosyltransferase and α 1,3-fucosyltransferase from *Helicobacter pylori* has also been characterized (Martin *et al.*, *J. Biol. Chem.* **272**: 21349-21356 (1997)). Also of use in the present invention are the glycosyltransferases of *Campylobacter jejuni* (*see*, for example, http://afmb.cnrs-mrs.fr/~pedro/CAZY/gtf_42.html).

10 a) Fucosyltransferases

[0364] In some embodiments, a glycosyltransferase used in the method of the invention is a fucosyltransferase. Fucosyltransferases are known to those of skill in the art. Exemplary fucosyltransferases include enzymes, which transfer L-fucose from GDP-fucose to a hydroxy position of an acceptor sugar. Fucosyltransferases that transfer non-nucleotide sugars to an acceptor are also of use in the present invention.

[0365] In some embodiments, the acceptor sugar is, for example, the GlcNAc in a Gal β (1 \rightarrow 3,4)GlcNAc β - group in an oligosaccharide glycoside. Suitable fucosyltransferases for this reaction include the Gal β (1 \rightarrow 3,4)GlcNAc β 1- α (1 \rightarrow 3,4)fucosyltransferase (FTIII E.C. No. 2.4.1.65), which was first characterized from human milk (*see*, Palcic, *et al.*, *Carbohydrate Res.* **190**: 1-11 (1989); Prieels, *et al.*, *J. Biol. Chem.* **256**: 10456-10463 (1981); and Nunez, *et al.*, *Can. J. Chem.* **59**: 2086-2095 (1981)) and the Gal β (1 \rightarrow 4)GlcNAc β - α fucosyltransferases (FTIV, FTV, FTVI) which are found in human serum. FTVII (E.C. No. 2.4.1.65), a sialyl α (2 \rightarrow 3)Gal β ((1 \rightarrow 3)GlcNAc β fucosyltransferase, has also been characterized. A recombinant form of the Gal β (1 \rightarrow 3,4) GlcNAc β -

α (1 \rightarrow 3,4)fucosyltransferase has also been characterized (*see*, Dumas, *et al.*, *Bioorg. Med. Letters* **1**: 425-428 (1991) and Kukowska-Latallo, *et al.*, *Genes and Development* **4**: 1288-1303 (1990)). Other exemplary fucosyltransferases include, for example, α 1,2 fucosyltransferase (E.C. No. 2.4.1.69). Enzymatic fucosylation can be carried out by the methods described in Mollicone, *et al.*, *Eur. J. Biochem.* **191**: 169-176 (1990) or U.S. Patent No. 5,374,655. Cells that are used to produce a fucosyltransferase will also include an enzymatic system for synthesizing GDP-fucose.

b) *Galactosyltransferases*

[0366] In another group of embodiments, the glycosyltransferase is a galactosyltransferase. Exemplary galactosyltransferases include $\alpha(1,3)$ galactosyltransferases (E.C. No. 2.4.1.151, see, e.g., Dabkowski *et al.*, *Transplant Proc.* **25**:2921 (1993) and Yamamoto *et al.* *Nature* **345**: 229-233 (1990), bovine (GenBank j04989, Joziase *et al.*, *J. Biol. Chem.* **264**: 14290-14297 (1989)), murine (GenBank m26925; Larsen *et al.*, *Proc. Nat'l. Acad. Sci. USA* **86**: 8227-8231 (1989)), porcine (GenBank L36152; Strahan *et al.*, *Immunogenetics* **41**: 101-105 (1995)). Another suitable $\alpha(1,3)$ galactosyltransferase is that which is involved in synthesis of the blood group B antigen (EC 2.4.1.37, Yamamoto *et al.*, *J. Biol. Chem.* **265**: 1146-1151 (1990) (human)). Yet a further exemplary galactosyltransferase is core Gal-T1.

[0367] Also suitable for use in the methods of the invention are $\beta(1,4)$ galactosyltransferases, which include, for example, EC 2.4.1.90 (LacNAc synthetase) and EC 2.4.1.22 (lactose synthetase) (bovine (D'Agostaro *et al.*, *Eur. J. Biochem.* **183**: 211-217 (1989)), human (Masri *et al.*, *Biochem. Biophys. Res. Commun.* **157**: 657-663 (1988)), murine (Nakazawa *et al.*, *J. Biochem.* **104**: 165-168 (1988)), as well as E.C. 2.4.1.38 and the ceramide galactosyltransferase (EC 2.4.1.45, Stahl *et al.*, *J. Neurosci. Res.* **38**: 234-242 (1994)). Other suitable galactosyltransferases include, for example, $\alpha(1,2)$ galactosyltransferases (from e.g., *Schizosaccharomyces pombe*, Chapell *et al.*, *Mol. Biol. Cell* **5**: 519-528 (1994)).

c) *Sialyltransferases*

[0368] Sialyltransferases are another type of glycosyltransferase that is useful in the recombinant cells and reaction mixtures of the invention. Cells that produce recombinant sialyltransferases will also produce CMP-sialic acid, which is a sialic acid donor for sialyltransferases. Examples of sialyltransferases that are suitable for use in the present invention include ST3Gal III (e.g., a rat or human ST3Gal III), ST3Gal IV, ST3Gal I, ST6Gal I, ST3Gal V, ST6Gal II, ST6GalNAc I, ST6GalNAc II, and ST6GalNAc III (the sialyltransferase nomenclature used herein is as described in Tsuji *et al.*, *Glycobiology* **6**: v-xiv (1996)). An exemplary $\alpha(2,3)$ sialyltransferase referred to as $\alpha(2,3)$ sialyltransferase (EC 2.4.99.6) transfers sialic acid to the non-reducing terminal Gal of a Gal β 1 \rightarrow 3Glc disaccharide or glycoside. See, Van den Eijnden *et al.*, *J. Biol. Chem.* **256**: 3159 (1981), Weinstein *et al.*, *J. Biol. Chem.* **257**: 13845 (1982) and Wen *et al.*, *J. Biol. Chem.* **267**: 21011 (1992). Another exemplary $\alpha(2,3)$ -sialyltransferase (EC 2.4.99.4) transfers sialic acid to the non-reducing

terminal Gal of the disaccharide or glycoside. *see*, Rearick *et al.*, *J. Biol. Chem.* **254**: 4444 (1979) and Gillespie *et al.*, *J. Biol. Chem.* **267**: 21004 (1992). Further exemplary enzymes include Gal- β -1,4-GlcNAc α -2,6 sialyltransferase (*See*, Kurosawa *et al.* *Eur. J. Biochem.* **219**: 375-381 (1994)).

- 5 **[0369]** Preferably, for glycosylation of carbohydrates of glycopeptides the sialyltransferase will be able to transfer sialic acid to the sequence Gal β 1,4GlcNAc-, the most common penultimate sequence underlying the terminal sialic acid on fully sialylated carbohydrate structures (*see*, Table 5).

Table 5: Sialyltransferases which use the Gal β 1,4GlcNAc sequence as an acceptor substrate

Sialyltransferase	Source	Sequence(s) formed	Ref.
ST6Gal I	Mammalian	NeuAcI2,6Gal β 1,4GlcNAc-	1
ST3Gal III	Mammalian	NeuAcI2,3Gal β 1,4GlcNAc- NeuAcI2,3Gal β 1,3GlcNAc-	1
ST3Gal IV	Mammalian	NeuAcI2,3Gal β 1,4GlcNAc- NeuAcI2,3Gal β 1,3GlcNAc-	1
ST6Gal II	Mammalian	NeuAcI2,6Gal β 1,4GlcNA	
ST6Gal II	photobacterium	NeuAcI2,6Gal β 1,4GlcNAc-	2
ST3Gal V	<i>N. meningitides</i> <i>N. gonorrhoeae</i>	NeuAcI2,3Gal β 1,4GlcNAc-	3

- 1) Goochee *et al.*, *Bio/Technology* **9**: 1347-1355 (1991)
 2) Yamamoto *et al.*, *J. Biochem.* **120**: 104-110 (1996)
 3) Gilbert *et al.*, *J. Biol. Chem.* **271**: 28271-28276 (1996)

- [0370]** An example of a sialyltransferase that is useful in the claimed methods is ST3Gal
 15 III, which is also referred to as α (2,3)sialyltransferase (EC 2.4.99.6). This enzyme catalyzes the transfer of sialic acid to the Gal of a Gal β 1,3GlcNAc or Gal β 1,4GlcNAc glycoside (*see*, *e.g.*, Wen *et al.*, *J. Biol. Chem.* **267**: 21011 (1992); Van den Eijnden *et al.*, *J. Biol. Chem.* **256**: 3159 (1991)) and is responsible for sialylation of asparagine-linked oligosaccharides in glycopeptides. The sialic acid is linked to a Gal with the formation of an α -linkage between
 20 the two saccharides. Bonding (linkage) between the saccharides is between the 2-position of NeuAc and the 3-position of Gal. This particular enzyme can be isolated from rat liver (Weinstein *et al.*, *J. Biol. Chem.* **257**: 13845 (1982)); the human cDNA (Sasaki *et al.* (1993) *J. Biol. Chem.* **268**: 22782-22787; Kitagawa & Paulson (1994) *J. Biol. Chem.* **269**: 1394-

1401) and genomic (Kitagawa *et al.* (1996) *J. Biol. Chem.* **271**: 931-938) DNA sequences are known, facilitating production of this enzyme by recombinant expression. In a preferred embodiment, the claimed sialylation methods use a rat ST3Gal III.

[0371] Other exemplary sialyltransferases of use in the present invention include those isolated from *Campylobacter jejuni*, including the $\alpha(2,3)$. See, e.g., WO99/49051.

[0372] Sialyltransferases other than those listed in Table 5, are also useful in an economic and efficient large-scale process for sialylation of commercially important glycopeptides. As a simple test to find out the utility of these other enzymes, various amounts of each enzyme (1-100 mU/mg protein) are reacted with asialo- α_1 AGP (at 1-10 mg/ml) to compare the ability of the sialyltransferase of interest to sialylate glycopeptides relative to either bovine ST6Gal I, ST3Gal III or both sialyltransferases. Alternatively, other glycopeptides, or N-linked oligosaccharides enzymatically released from the peptide backbone can be used in place of asialo- α_1 AGP for this evaluation. Sialyltransferases with the ability to sialylate N-linked oligosaccharides of glycopeptides more efficiently than ST6Gal I are useful in a practical large-scale process for peptide sialylation (as illustrated for ST3Gal III in this disclosure).

d) GalNAc transferases

[0373] N-acetylgalactosaminyltransferases are of use in practicing the present invention, particularly for binding a GalNAc moiety to an amino acid of the O-linked glycosylation site of the peptide. Suitable N-acetylgalactosaminyltransferases include, but are not limited to, $\alpha(1,3)$ N-acetylgalactosaminyltransferase, $\beta(1,4)$ N-acetylgalactosaminyltransferases (Nagata *et al.*, *J. Biol. Chem.* **267**: 12082-12089 (1992) and Smith *et al.*, *J. Biol. Chem.* **269**: 15162 (1994)) and polypeptide N-acetylgalactosaminyltransferase (Homa *et al.*, *J. Biol. Chem.* **268**: 12609 (1993)).

[0374] Production of proteins such as the enzyme GalNAc T_{I-XX} from cloned genes by genetic engineering is well known. See, eg., U.S. Pat. No. 4,761,371. One method involves collection of sufficient samples, then the amino acid sequence of the enzyme is determined by N-terminal sequencing. This information is then used to isolate a cDNA clone encoding a full-length (membrane bound) transferase which upon expression in the insect cell line Sf9 resulted in the synthesis of a fully active enzyme. The acceptor specificity of the enzyme is then determined using a semiquantitative analysis of the amino acids surrounding known

glycosylation sites in 16 different proteins followed by in vitro glycosylation studies of synthetic peptides. This work has demonstrated that certain amino acid residues are overrepresented in glycosylated peptide segments and that residues in specific positions surrounding glycosylated serine and threonine residues may have a more marked influence on acceptor efficiency than other amino acid moieties.

2. *Sulfotransferases*

[0375] The invention also provides methods for producing peptides that include sulfated molecules, including, for example sulfated polysaccharides such as heparin, heparan sulfate, carragenen, and related compounds. Suitable sulfotransferases include, for example, chondroitin-6-sulphotransferase (chicken cDNA described by Fukuta *et al.*, *J. Biol. Chem.* **270**: 18575-18580 (1995); GenBank Accession No. D49915), glycosaminoglycan N-acetylglucosamine N-deacetylase/N-sulphotransferase 1 (Dixon *et al.*, *Genomics* **26**: 239-241 (1995); UL18918), and glycosaminoglycan N-acetylglucosamine N-deacetylase/N-sulphotransferase 2 (murine cDNA described in Orellana *et al.*, *J. Biol. Chem.* **269**: 2270-2276 (1994) and Eriksson *et al.*, *J. Biol. Chem.* **269**: 10438-10443 (1994); human cDNA described in GenBank Accession No. U2304).

3. *Cell-Bound Glycosyltransferases*

[0376] In another embodiment, the enzymes utilized in the method of the invention are cell-bound glycosyltransferases. Although many soluble glycosyltransferases are known (see, for example, U.S. Pat. No. 5,032,519), glycosyltransferases are generally in membrane-bound form when associated with cells. Many of the membrane-bound enzymes studied thus far are considered to be intrinsic proteins; that is, they are not released from the membranes by sonication and require detergents for solubilization. Surface glycosyltransferases have been identified on the surfaces of vertebrate and invertebrate cells, and it has also been recognized that these surface transferases maintain catalytic activity under physiological conditions. However, the more recognized function of cell surface glycosyltransferases is for intercellular recognition (Roth, *MOLECULAR APPROACHES TO SUPRACELLULAR PHENOMENA*, 1990).

[0377] Methods have been developed to alter the glycosyltransferases expressed by cells. For example, Larsen *et al.*, *Proc. Natl. Acad. Sci. USA* **86**: 8227-8231 (1989), report a genetic approach to isolate cloned cDNA sequences that determine expression of cell surface oligosaccharide structures and their cognate glycosyltransferases. A cDNA library generated

from mRNA isolated from a murine cell line known to express UDP-galactose:β-D-galactosyl-1,4-N-acetyl-D-glucosaminide α-1,3-galactosyltransferase was transfected into COS-1 cells. The transfected cells were then cultured and assayed for α 1-3 galactosyltransferase activity.

- 5 [0378] Francisco *et al.*, *Proc. Natl. Acad. Sci. USA* **89**: 2713-2717 (1992), disclose a method of anchoring β-lactamase to the external surface of *Escherichia coli*. A tripartite fusion consisting of (i) a signal sequence of an outer membrane protein, (ii) a membrane-spanning section of an outer membrane protein, and (iii) a complete mature β-lactamase sequence is produced resulting in an active surface bound β-lactamase molecule. However,
10 the Francisco method is limited only to procaryotic cell systems and as recognized by the authors, requires the complete tripartite fusion for proper functioning.

4. Fusion Proteins

- [0379] In other exemplary embodiments, the methods of the invention utilize fusion proteins that have more than one enzymatic activity that is involved in synthesis of a desired
15 glycopeptide conjugate. The fusion polypeptides can be composed of, for example, a catalytically active domain of a glycosyltransferase that is joined to a catalytically active domain of an accessory enzyme. The accessory enzyme catalytic domain can, for example, catalyze a step in the formation of a nucleotide sugar that is a donor for the glycosyltransferase, or catalyze a reaction involved in a glycosyltransferase cycle. For
20 example, a polynucleotide that encodes a glycosyltransferase can be joined, in-frame, to a polynucleotide that encodes an enzyme involved in nucleotide sugar synthesis. The resulting fusion protein can then catalyze not only the synthesis of the nucleotide sugar, but also the transfer of the sugar moiety to the acceptor molecule. The fusion protein can be two or more cycle enzymes linked into one expressible nucleotide sequence. In other embodiments the
25 fusion protein includes the catalytically active domains of two or more glycosyltransferases. See, for example, 5,641,668. The modified glycopeptides of the present invention can be readily designed and manufactured utilizing various suitable fusion proteins (*see*, for example, PCT Patent Application PCT/CA98/01180, which was published as WO 99/31224 on June 24, 1999.)

5. *Immobilized Enzymes*

[0380] In addition to cell-bound enzymes, the present invention also provides for the use of enzymes that are immobilized on a solid and/or soluble support. In an exemplary embodiment, there is provided a glycosyltransferase that is conjugated to a PEG via an intact glycosyl linker according to the methods of the invention. The PEG-linker-enzyme conjugate is optionally attached to solid support. The use of solid supported enzymes in the methods of the invention simplifies the work up of the reaction mixture and purification of the reaction product, and also enables the facile recovery of the enzyme. The glycosyltransferase conjugate is utilized in the methods of the invention. Other combinations of enzymes and supports will be apparent to those of skill in the art.

Purification of Peptide Conjugates

[0381] The products produced by the above processes can be used without purification. However, it is usually preferred to recover the product. Standard, well-known techniques for recovery of glycosylated saccharides such as thin or thick layer chromatography, column chromatography, ion exchange chromatography, or membrane filtration can be used. It is preferred to use membrane filtration, more preferably utilizing a reverse osmotic membrane, or one or more column chromatographic techniques for the recovery as is discussed hereinafter and in the literature cited herein. For instance, membrane filtration wherein the membranes have molecular weight cutoff of about 3000 to about 10,000 can be used to remove proteins such as glycosyl transferases. Nanofiltration or reverse osmosis can then be used to remove salts and/or purify the product saccharides (*see, e.g.*, WO 98/15581). Nanofilter membranes are a class of reverse osmosis membranes that pass monovalent salts but retain polyvalent salts and uncharged solutes larger than about 100 to about 2,000 Daltons, depending upon the membrane used. Thus, in a typical application, saccharides prepared by the methods of the present invention will be retained in the membrane and contaminating salts will pass through.

[0382] If the modified glycoprotein is produced intracellularly, as a first step, the particulate debris, either host cells or lysed fragments, is removed, for example, by centrifugation or ultrafiltration; optionally, the protein may be concentrated with a commercially available protein concentration filter, followed by separating the polypeptide variant from other impurities by one or more steps selected from immunoaffinity chromatography, ion-exchange column fractionation (*e.g.*, on diethylaminoethyl (DEAE) or

matrices containing carboxymethyl or sulfopropyl groups), chromatography on Blue-Sepharose, CM Blue-Sepharose, MONO-Q, MONO-S, lentil lectin-Sepharose, WGA-Sepharose, Con A-Sepharose, Ether Toyopearl, Butyl Toyopearl, Phenyl Toyopearl, SP-Sepharose, or protein A Sepharose, SDS-PAGE chromatography, silica chromatography, chromatofocusing, reverse phase HPLC (*e.g.*, silica gel with appended aliphatic groups), gel filtration using, *e.g.*, Sephadex molecular sieve or size-exclusion chromatography, chromatography on columns that selectively bind the polypeptide, and ethanol or ammonium sulfate precipitation.

[0383] Modified glycopeptides produced in culture are usually isolated by initial extraction from cells, enzymes, etc., followed by one or more concentration, salting-out, aqueous ion-exchange, or size-exclusion chromatography steps, *e.g.*, SP Sepharose. Additionally, the modified glycoprotein may be purified by affinity chromatography. HPLC may also be employed for one or more purification steps.

[0384] A protease inhibitor, *e.g.*, methylsulfonylfluoride (PMSF) may be included in any of the foregoing steps to inhibit proteolysis and antibiotics may be included to prevent the growth of adventitious contaminants.

[0385] Within another embodiment, supernatants from systems which produce the modified glycopeptide of the invention are first concentrated using a commercially available protein concentration filter, for example, an Amicon or Millipore Pellicon ultrafiltration unit. Following the concentration step, the concentrate may be applied to a suitable purification matrix. For example, a suitable affinity matrix may comprise a ligand for the peptide, a lectin or antibody molecule bound to a suitable support. Alternatively, an anion-exchange resin may be employed, for example, a matrix or substrate having pendant DEAE groups. Suitable matrices include acrylamide, agarose, dextran, cellulose, or other types commonly employed in protein purification. Alternatively, a cation-exchange step may be employed. Suitable cation exchangers include various insoluble matrices comprising sulfopropyl or carboxymethyl groups. Sulfopropyl groups are particularly preferred.

[0386] Finally, one or more RP-HPLC steps employing hydrophobic RP-HPLC media, *e.g.*, silica gel having pendant methyl or other aliphatic groups, may be employed to further purify a polypeptide variant composition. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a homogeneous modified glycoprotein.

[0387] The modified glycopeptide of the invention resulting from a large-scale fermentation may be purified by methods analogous to those disclosed by Urdal *et al.*, *J. Chromatog.* **296**: 171 (1984). This reference describes two sequential, RP-HPLC steps for purification of recombinant human IL-2 on a preparative HPLC column. Alternatively, techniques such as affinity chromatography may be utilized to purify the modified glycoprotein.

Pharmaceutical Compositions

[0388] Polypeptides modified at various O-linked glycosylation site according to the method of the present invention have a broad range of pharmaceutical applications. For example, GLP-1 may be used for the treatment or prevention of diabetes or obesity.

[0389] An additional example, human growth hormone (hGH) modified according to the methods of the present invention may be used to treat growth-related conditions such as dwarfism, short-stature in children and adults, cachexia/muscle wasting, general muscular atrophy, and sex chromosome abnormality (*e.g.*, Turner's Syndrome). Other conditions may be treated using modified hGH include: short-bowel syndrome, lipodystrophy, osteoporosis, uraemia, burns, female infertility, bone regeneration, general diabetes, type II diabetes, osteo-arthritis, chronic obstructive pulmonary disease (COPD), and insomnia. Moreover, modified hGH may also be used to promote various processes, *e.g.*, general tissue regeneration, bone regeneration, and wound healing, or as a vaccine adjunct.

[0390] Thus, in one aspect, the invention provides a pharmaceutical composition. The pharmaceutical composition includes a pharmaceutically acceptable diluent and a covalent conjugate between a non-naturally-occurring, water-soluble polymer, therapeutic moiety or biomolecule and a glycosylated or non-glycosylated peptide. The polymer, therapeutic moiety or biomolecule is conjugated to the peptide via an intact glycosyl linking group interposed between and covalently linked to both the peptide and the polymer, therapeutic moiety or biomolecule.

[0391] Pharmaceutical compositions of the invention are suitable for use in a variety of drug delivery systems. Suitable formulations for use in the present invention are found in *Remington's Pharmaceutical Sciences*, Mace Publishing Company, Philadelphia, PA, 17th ed. (1985). For a brief review of methods for drug delivery, *see*, Langer, *Science* **249**:1527-1533 (1990).

[0392] The pharmaceutical compositions may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, intravenous, intracranial, intraperitoneal, subcutaneous or intramuscular administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable matrices, such as microspheres (*e.g.*, polylactate polyglycolate), may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268 and 5,075,109.

[0393] Commonly, the pharmaceutical compositions are administered subcutaneously or parenterally, *e.g.*, intravenously. Thus, the invention provides compositions for parenteral administration which comprise the compound dissolved or suspended in an acceptable carrier, preferably an aqueous carrier, *e.g.*, water, buffered water, saline, PBS and the like.

The compositions may also contain detergents such as Tween 20 and Tween 80; stabilizers such as mannitol, sorbitol, sucrose, and trehalose; and preservatives such as EDTA and m-cresol. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as pH adjusting and buffering agents, tonicity adjusting agents, wetting agents, detergents and the like.

[0394] These compositions may be sterilized by conventional sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile aqueous carrier prior to administration. The pH of the preparations typically will be between 3 and 11, more preferably from 5 to 9 and most preferably from 7 and 8.

[0395] In some embodiments the glycopeptides of the invention can be incorporated into liposomes formed from standard vesicle-forming lipids. A variety of methods are available for preparing liposomes, as described in, *e.g.*, Szoka *et al.*, *Ann. Rev. Biophys. Bioeng.* **9**: 467 (1980), U.S. Pat. Nos. 4,235,871, 4,501,728 and 4,837,028. The targeting of liposomes using a variety of targeting agents (*e.g.*, the sialyl galactosides of the invention) is well known in the art (*see, e.g.*, U.S. Patent Nos. 4,957,773 and 4,603,044).

[0396] Standard methods for coupling targeting agents to liposomes can be used. These methods generally involve incorporation into liposomes of lipid components, such as

phosphatidylethanolamine, which can be activated for attachment of targeting agents, or derivatized lipophilic compounds, such as lipid-derivatized glycopeptides of the invention.

[0397] Targeting mechanisms generally require that the targeting agents be positioned on the surface of the liposome in such a manner that the target moieties are available for

5 interaction with the target, for example, a cell surface receptor. The carbohydrates of the invention may be attached to a lipid molecule before the liposome is formed using methods known to those of skill in the art (*e.g.*, alkylation or acylation of a hydroxyl group present on the carbohydrate with a long chain alkyl halide or with a fatty acid, respectively).

Alternatively, the liposome may be fashioned in such a way that a connector portion is first
10 incorporated into the membrane at the time of forming the membrane. The connector portion must have a lipophilic portion, which is firmly embedded and anchored in the membrane. It must also have a reactive portion, which is chemically available on the aqueous surface of the liposome. The reactive portion is selected so that it will be chemically suitable to form a stable chemical bond with the targeting agent or carbohydrate, which is added later. In some
15 cases it is possible to attach the target agent to the connector molecule directly, but in most instances it is more suitable to use a third molecule to act as a chemical bridge, thus linking the connector molecule which is in the membrane with the target agent or carbohydrate which is extended, three dimensionally, off of the vesicle surface.

[0398] The following examples are provided to illustrate the conjugates, and methods and
20 of the present invention, but not to limit the claimed invention.

EXAMPLES

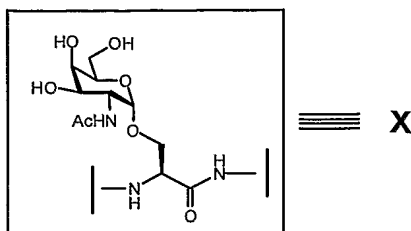
EXAMPLE 1

1.1 Preparation of Glucagon-Like Peptide Mutants Comprising Artificial Glycosylation Sites

25 [0399] Mutations in the amino acid sequence of Glucagon-Like Peptide-1 (GLP-1) will be made in order to introduce sites for O-linked glycosylation, such that the protein may be modified at these sites using the method of the present invention. Mutatants can be created using well known methods for solid state synthesis. Alternatively, mutations will be introduced into a nucleic acid the sequence encoding GLP-1 such that O-linked glycosylation
30 sites will be introduced at each position along the peptide back bone.

[0400] The following are some exemplary GLP-1 mutants.

GLP-1 Glycopeptides



Ac-X-HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 H-X-EGTFTSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HA-X-GTFTSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAE-X-TFTSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEG-X-FTSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGT-X-TSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTF-X-SDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFT-X-DVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTS-X-VSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSD-X-SSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDV-X-SYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVS-X-YLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSS-X-LEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSY-X-EGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYL-X-GQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLE-X-QAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEG-X-AAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQ-X-AKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQA-X-KEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAA-X-EFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAK-X-FIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKE-X-IAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEF-X-AWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFI-X-WLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIA-X-LVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIW-X-VKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAWL-X-KGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAWLX-X-GR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAWLXK-X-R-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAWLXKG-X-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAWLXKGR-X-NH₂

1.2 Preparation of GLP-1-GalNAc (pH 6.2)

[0401] GLP-1 (960 µg) in 3.2 mL of buffer will be concentrated by ultrafiltration using an
 5 UF filter (5 KDa) and then reconstituted with 1 mL of 25 mM MES buffer (pH 6.2, 0.005%
 NaN₃). The UDP-GalNAc (6 mg, 9.24 mM), GalNAc-T2 (40 µL, 0.04 U), and 100 mM
 MnCl₂ (40 µL, 4 mM) will then be added and the resulting solution will be incubated at room

temperature. After 48 h, the MALDI should the reaction was complete (shift of the mass ion from 18800 to 19023 mass units). The reaction mixture will be purified by HPLC using SEC (Superdex 75 and Superdex 200). The column will be eluted using phosphate buffered saline, pH 4.9 and 0.005% tween 80. The peak corresponding to GLP-1-GalNAc will be collected
 5 and concentrated to about 150 μ L using a Centricon 5 KDa filter and the volume will be adjusted to 1 mL using PBS (phosphate buffered saline, pH 4.9 and 0.005% tween 80); protein concentration 1 mg/mL A_{280}).

1.3 Preparation of GLP-1-GalNAc-Gal (pH 6.0)

[0402] GLP-1-GalNAc (100 μ g) will be added to a 100 μ L of a solution containing 25 mM
 10 MES buffer, pH 6.0, 1.5 mM UDP-GalNAc, 10 mM $MgCl_2$ and 80 mU GalNAc-T2. The CMP-SA-PEG-20 KDa (0.5 mg, 0.025 μ mole), UDP-galactose 75 μ g (0.125 μ mole), core-1-Gal-T 20 μ L (10 mU) will then be added and the solution slowly rocked at 32 °C for 24 h. MALDI should indicate complete conversion of GLP-1-GalNAc into GLP-1-GalNAc-Gal.

1.4 Preparation of GLP-1-F-GalNAc-SA-PEG-20 KDa (C).

1.3a Sequential Process (pH 6.2).

[0403] A GLP-1-GalNAc solution containing 1 mg of protein will be buffer exchanged into
 25 mM MES buffer (pH 6.2, 0.005% NaN_3) and CMP-SA-PEG (20KDa) (5 mg, 0.25 μ mole). $MnCl_2$ (100 μ L, 100 mM solution) and ST6GalNAc-I (100 μ L) will be added and the
 reaction mixture will be rocked slowly at 32 °C. Aliquots will be taken at time points (24, 48
 20 and 72 h) and analyzed by SDS-PAGE. After 24 h, no further reaction should be observed. The reaction mixture will be concentrated by spin filtration (5 KDa), buffer exchanged with 25 mM NaOAc (pH 4.9) and concentrated to 1 mL. The product will be purified using ion exchange (SP-Sepharose, 25 mM NaOAc, pH 4.9) and SEC (Superdex 75; PBS-pH 7.2, 0.005% tween 80, 1 ml/min). The desired fraction will be collected, concentrated to 0.5 mL
 25 and stored at 4 °C.

1.4b One Pot process using ST6GalNAc-I (pH 6.0)

[0404] GLP-1 (960 μ g of protein dissolved in 3.2 mL of product formulation buffer) will be
 concentrated by spin filtration (5 KDa) to 0.5 mL and reconstituted in 25 mM MES buffer
 (pH 6.0, 0.005% NaN_3) to a total volume of about 1 mL, or a protein concentration of 1
 30 mg/mL. UDP-GalNAc (6 mg, 9.21 μ mol), GalNAc-T2 (80 μ L, 80 mU), CMP-SA-PEG (20 KDa) (6 mg, 0.3 μ mol) and mouse enzyme ST6GalNAc-I (120 μ L will be added). The

solution will be rocked at 32 °C for 48 h and purified using standard chromatography conditions on SP-Sepharose and SEC as described above. A total of 0.5 mg of protein (A₂₈₀) should be obtained, or about a 50% overall yield. The product structure will be confirmed by analysis with both MALDI and SDS-PAGE.

5 **1.5 Preparation of GLP-1-GalNAc-Gal-SA-PEG-20 KDa (D)**

1.5a Starting from GLP-1-GalNAc

[0405] UDP-galactose (4 mg, 6.5 µmole), core-1-Gal-T₁ (320 µL, 160 mU), CMP-SA-PEG-20 KDa (8 mg, 0.4 µmole), ST3Gal2 (80 µL, 0.07 mU) and 100 mM MnCl₂ (80 µL) will be directly added to the crude reaction mixture of the GLP-1-GalNAc (1.5 mg) in 25 mM
10 MES buffer (pH 6.0), 1.5 mL, as described above. The resulting mixture will be incubated at 32 °C for 60 h, however, the reaction should be complete after 24 h. The reaction mixture will be centrifuged and the solution was concentrated to 0.2 mL using ultrafiltration (5 KDa) and then redissolved in 25 mM NaOAc (pH 4.5) to a final volume of 1 mL. The product will be purified using SP-Sepharose, the peak fractions were concentrated using a spin filter
15 (5KDa), and the residue purified further using SEC (Superdex 75). After concentration using a spin filter (5 KDa), the protein will be diluted to 1 mL using formulation buffer (PBS, 2.5% mannitol, 0.005% polysorbate, pH 6.5) and formulated at a protein concentration of 850 µg protein per mL (A₂₈₀). The overall yield should be around 55%.

1.5b Starting from GLP-1

20 [0406] GLP-1 (960 µg, 3.2 mL) will be concentrated by spin filter (5 KDa) and reconstituted with 25 mM MES buffer (pH 6.0, 0.005% NaN₃). The total volume of the GLP-1 solution will be adjusted to about 1 mg/mL and UDP-GalNAc (6 mg), GalNAc-T2 (80 µL), UDP-galactose (6 mg), core-1-Gal-T₁ (160 µL, 80 µU), CMP-SA-PEG (20 KDa) (6 mg), ST3Gal-2 (160 µL, 120 µU) and MnCl₂ (40 µL of a 100 mM solution) will be added.
25 The resulting mixture will be incubated at 32 °C for 48 h.

1.6 SP Sepharose HPLC Chromatography.

[0407] The SP Sepharose column (HiTrap HP, FF, 1 mL, Amersham) can be used with a Varian HPLC system to separate individual GLP-1 peptides from crude extracts. Absorbance at 280 nm will be monitored. The column will be washed with 20 mL of 2 M NaCl in 25 mM
30 sodium acetate (pH 4.5) contained 0.005% polysorbate 80 and was equilibrated with 20 mL of 25 mM sodium acetate (pH 4.5) contained 0.005% polysorbate 80 at a flow rate of 1.0

mL/min. The sample (about 0.5 mg /200 µL) will be injected onto the column and the product will be eluted using the gradient: 0-10 min, 25 mM NaAc, pH 4.5, 0.005% Polysorbate 80; 10-20 min, a gradient of 0-0.5 M NaCl in 25 mM NaAc, pH 4.5, 0.005% Polysorbate 80; 20-25 min, a gradient of 0.5 M-0.0 M NaCl in 25 mM NaAc, pH 4.5, 0.005% Polysorbate 80; and 25-30 min, 25 mM NaAc, pH 4.5, 0.005% Polysorbate 80). Fractions will be collected and concentrated to about 1 mL by using 5 KDa filter for analysis and further purification. Samples will be stored at 4 °C.

1.6 Size Exclusion Chromatography

[0408] A Varian HPLC system containing a Superdex 75 column (HR 10/30, 10 x 300 mm, Amersham) will be used at a flow rate of 1.0 mL/min, while monitoring absorbance at 280 nm. The sample will be injected (about 0.2 mg/200 µL) and eluted with PBS, pH 7.4, 0.005% Polysorbate 80. Fractions will be collected and concentrated to about 1 mL by using a 5 KDa filter. Samples will be stored at 4 °C.

1.7 SDS PAGE Analysis

[0409] 4-20% acrylamide gradient slab gels will be used. Samples will be mixed with SDS Sample Buffer contained 1 mM DTT, and heated at 85 °C for 6 min. Samples will be run in gel under a consistent voltage at 125 mV for 1 h 50 min. After electrophoresis, the proteins will be stained with colloidal stain solution at room temperature for 2-24 hours depended on the protein concentration. The standard proteins shown on Tris-Glycine gel will be myosin (250 KDa), phosphorylase (148KDa), BSA (98 KDa, glutamic dehydrogenase (64 KDa), alcohol dehydrogenase (50 KDa), carbonic anhydrase (36 KDa), lysozyme (22KDa), aprotinin (6 KDa), and insulin β-chain (4 KDa). The protein bands in wet gel will be visualized using an HP Scanjet 7400C, and the picture of gel will be optimized using the HP Precision Scan Program.

1.8 MALDI Analysis

[0411] Samples will be dialyzed for 45 min using an MF-Millipore membrane filter (0.025 µm pore, 47 mm dia), floating on water. The dialyzed aliquots will be dried on a speedvac, re-dissolved in a small amount of water, and mixed with a solution of 2,5-dihydroxybenzoic acid (9 g/L) and 5- methoxysalicylic acid (1 g/L) dissolved in water/acetonitrile (50:50). The mixtures will be dried onto the MALDI target and analyzed using an Applied Biosystems

DE-Pro mass spectrometer operated in the linear/negative-ion mode (Analytic lab, Neose Tech., Horsham, PA).

1.9 Peptide Mapping Analysis

[0412] Protein sample will be digested by trypsin overnight at 37 °C and loaded on a LC-MS system equipped with a Finnigan LCQ-classic ion trap mass spectrometer system with a
5 electrospray ion source interfaced to a 15 cm x 300 um id LC Packings PepMap reverse-phase capillary chromatography column. 1 µL volume of the extract will be injected and the peptides will be eluted from the column using a CH₃CN/0.1% formic acid gradient at a flow rate of 3 µL/min. The electrospray ion source will be operated at 4.0 kV. The digest will be
10 analyzed using the data dependent multitask capability of the instrument acquiring full scan mass spectra to determine peptide molecular weights and product ion spectra to determine amino acid sequence in successive instrument scans. This mode of analysis produced approximately 100 collisionally induced dissociation (CID) spectra of ions ranging in abundance over several orders of magnitude.

15 [0413] The data will be analyzed by locating the ten to fifteen most abundant ions in a base peak presentation of the full scan data and interpreting the CID spectra of those ions to produce the tabulated results for each digest.

1.10 Protein Concentration Assay

[0414] Protein concentration will be determined by spectrophotometer at a fixed
20 absorbance of 280 nm with 1 cm path length of cell. Triplicate readings will be measured for a tested sample with water and buffer as controls. Protein concentration will be determined using extinction coefficient at 0.799 mL/mg protein.

1.11 Formulation of Final Product

[0415] The formulation buffer contained pyrogen-free PBS, pH 6.5, 2.5% mannitol, and
25 0.05% Polysorbate 80 that will be degassed by vacuum and sterile filtered (0.2 µm).

[0416] Any endotoxin will be removed using a Detoxi-GelTM equilibrated with 5 column beds of the formulation buffer (PBS, pH 6.5, 2.5% mannitol, and 0.05% Polysorbate 80). The flow rate was controlled by gravity at ~ 0.3 mL/min. Product samples will be applied onto the gel, and the product eluted using the formulation buffer. The volume of the

collected product will be adjusted with additional formulation buffer to provide a protein concentration of about 100 µg/mL.

[0417] The peptide formulations will be sterile filtered (0.2 µ) and the effluent will be dispensed as 1 mL aliquots into 2.0 mL pyrogen-free vials. In addition, aliquots will be taken
5 for endotoxin and protein analysis. All products will be stored at 4 °C.

1.12 Endotoxin Determination

[0418] Endotoxin contamination will be determined using Limulus Amebocyte Lysate (LAL) assay (BioWhittaker, Kinetic-QCL Kit, Cat#: 50-650U).

EXAMPLE 2

10 Determination of Biological Activity of GLP-1 Peptides

[0419] Methods for measuring biological activity of GLP-1 are known in the art. In particular, GLP-1 activity will be measured *in vivo* and *in vitro* as disclosed in Xiao Q., *et al.* (2001) Biochemistry 40:2860-2869, which is incorporated herein by reference in its entirety.

15 [0420] While this invention has been disclosed with reference to specific embodiments, it is apparent that other embodiments and variations of this invention may be devised by others skilled in the art without departing from the true spirit and scope of the invention.

[0421] All patents, patent applications, and other publications cited in this application are incorporated by reference in the entirety.

WHAT IS CLAIMED IS:

1. A peptide having a formula selected from:



in which AA is an amino acid with a side chain that comprises a hydroxyl moiety; and
X a modifying group or it is a saccharyl moiety.

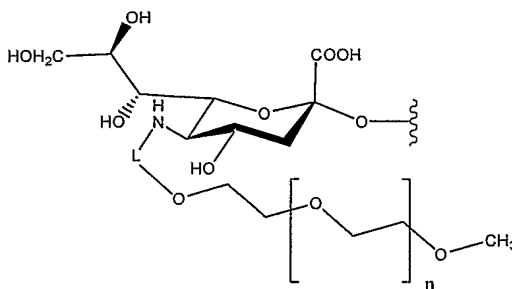
2. The peptide according to claim 1, wherein AA is introduced into said peptide via mutation of wild-type peptide.

3. The peptide according to claim 1, wherein X comprises a group selected from sialyl, galactosyl and Gal-Sia moieties, wherein at least one of said sialyl, galactosyl and Gal-Sia comprises a modifying group.

4. The peptide according to claim 1, wherein X comprises poly(ethylene glycol).

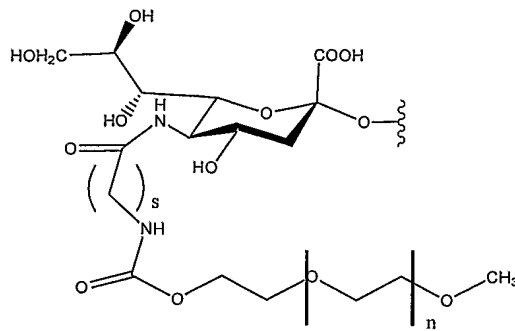
5. The peptide according to claim 1, wherein X comprises monomethoxy-poly(ethylene glycol).

6. The peptide according to claim 5, wherein X comprises the structure:



in which L is a substituted or unsubstituted alkyl or substituted or unsubstituted heteroalkyl group; and n is selected from the integers from 0 to about 500.

7. The peptide according to claim 5, wherein X comprises the structure:



in which s is selected from the integers from 0 to 20.

8. An isolated nucleic acid comprising a polynucleotide sequence encoding a mutant polypeptide, wherein the mutant polypeptide comprises an O-linked glycosylation site that does not exist in the corresponding wild-type polypeptide.

9. The nucleic acid of claim 8, wherein the polypeptide is a GLP-1 polypeptide.

10. An expression cassette comprising the nucleic acid of claim 8.

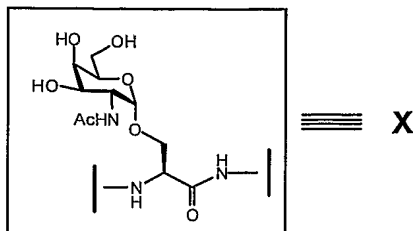
11. A cell comprising the nucleic acid of claim 8.

12. A method for making a glycoconjugate of a mutant polypeptide, which comprises an O-linked glycosylation that does not exist in the corresponding wild-type polypeptide, comprising the steps of:

- (a) recombinantly producing the mutant polypeptide, and
- (b) enzymatically glycosylating the mutant polypeptide with a modified sugar at said O-linked glycosylation site.

13. The method of claim 12, wherein the corresponding mutant polypeptide has an amino acid sequence selected from the group consisting of:

GLP-1 Glycopeptides



Ac-X-HAEGTFTSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 H-X-EGTFTSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HA-X-GTFTSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAE-X-TFTSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEG-X-FTSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGT-X-TSDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTF-X-SDVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFT-X-DVSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTS-X-VSSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSD-X-SSYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDV-X-SYLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVS-X-YLEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSS-X-LEGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSY-X-EGQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYL-X-GQAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLE-X-QAAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEG-X-AAKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQ-X-AKEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQA-X-KEFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAA-X-EFIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAK-X-FIAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKE-X-IAWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEF-X-AWLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFI-X-WLVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIA-X-LVKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAW-X-VKGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAWL-X-KGR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAWLX-X-GR-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAWLX-X-R-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAWLX-X-NH₂
 HAEGTFTSDVSSYLEGQAAKEFIAWLX-X-NH₂

3

4

1

2

3

4

14. A pharmaceutical composition of a Glucagon-Like Peptide-1

comprising an effective amount of a mutant polypeptide, which comprises an O-linked

glycosylation site that does not exist in the corresponding wild-type Glucagon-Like Peptide-

1, wherein said peptide is glycoconjugated with a modified sugar .

1 **15.** The pharmaceutical composition according to claim **13**, wherein said
2 modified sugar is modified with a member selected from poly(ethylene glycol) and m-
3 poly(ethylene glycol).

1 **16.** A method of providing Glucagon-Like Peptide-1 therapy to a subject
2 in need of said therapy, said method comprising, administering to said subject an amount of
3 an O-linked glyco-PEG-ylated Glucagon-Like Peptide-1 sufficient to provide a therapeutic
4 effect.

1 **17.** The method according to claim **16**, wherein said O-linked glyco-PEG-
2 ylated Glucagon-Like Peptide-1 is glyco-PEG-ylated on an amino acid residue not present in
3 wild type Glucagon-Like Peptide-1.

1